(predictive) purposes to thereby treat an individual prophylactically. Accordingly, one aspect of the present invention relates to diagnostic assays for determining INTERCEPT 340, MANGO 003, MANGO 347, TANGO 272, TANGO 295, TANGO 354, or TANGO 378 protein and/or nucleic acid expression as well as INTERCEPT 340, MANGO 003, MANGO 347, TANGO 272, TANGO 295, TANGO 354, or TANGO 378 activity, in the context of a biological sample (*e.g.*, blood, serum, cells, tissue) to thereby determine whether an individual is afflicted with a disease or disorder, or is at risk of developing a disorder, associated with aberrant or unwanted INTERCEPT 340, MANGO 003, MANGO 347, TANGO 272, TANGO 295, TANGO 354, or TANGO 378 gene expression or activity. The invention also provides for prognostic (or predictive) assays for determining whether an individual is at risk of developing a disorder associated with INTERCEPT 340, MANGO 003, MANGO 347, TANGO 272, TANGO 295, TANGO 354, or TANGO 378 protein or nucleic acid expression or activity. For example, mutations in a gene can be assayed in a biological sample. Such assays can be used for prognostic or predictive purpose to thereby prophylactically treat an individual prior to the onset of a disorder characterized by or associated with protein or nucleic acid expression or activity.

As an alternative to making determinations based on the absolute expression level of selected genes, determinations may be based on the normalized expression levels of these genes. Expression levels are normalized by correcting the absolute expression level of a INTERCEPT 340, MANGO 003, MANGO 347, TANGO 272, TANGO 295, TANGO 354, or TANGO 378 gene by comparing its expression to the expression of a gene that is not a INTERCEPT 340, MANGO 003, MANGO 347, TANGO 272, TANGO 295, TANGO 354, or TANGO 378, *e.g.*, a housekeeping gene that is constitutively expressed. Suitable genes for normalization include housekeeping genes such as the actin gene. This normalization allows the comparison of the expression level in one sample, *e.g.*, a patient sample, to another sample, *e.g.*, a non-disease sample, or between samples from different sources.

Alternatively, the expression level can be provided as a relative expression level. To determine a relative expression level of a gene, the level of expression of the gene is determined for 10 or more samples of different cell isolates, preferably 50 or more samples, prior to the determination of the expression level for the sample in question. The mean expression level of each of the genes assayed in the larger number of samples is determined and this is used as a baseline expression level for the gene(s) in question. The expression level of the gene determined for the test sample (absolute level of expression) is then divided by the mean expression value obtained for that gene. This provides a relative expression level and aids in identifying extreme cases of disease.

Preferably, the samples used in the baseline determination will be from diseased or from non-diseased cells of tissue. The choice of the cell source is dependent on the use of

the relative expression level. Using expression found in normal tissues as a mean expression score aids in validating whether the INTERCEPT 340, MANGO 003, MANGO 347, TANGO 272, TANGO 295, TANGO 354, or TANGO 378 gene assayed is diseased cell-type specific (versus normal cells). Such a use is particularly important in identifying whether a INTERCEPT 340, MANGO 003, MANGO 347, TANGO 272, TANGO 295, TANGO 354, or TANGO 378 gene can serve as a target gene. In addition, as more data is accumulated, the mean expression value can be revised, providing improved relative expression values based on accumulated data. Expression data from cells provide a means for grading the severity of the disease state.

Another aspect of the invention pertains to monitoring the influence of agents (*e.g.*, drugs, compounds) on the expression or activity of INTERCEPT 340, MANGO 003, MANGO 347, TANGO 272, TANGO 295, TANGO 354, or TANGO 378 genes in clinical trials.

These and other agents are described in further detail in the following sections.

1. Diagnostic Assays

An exemplary method for detecting the presence or absence of a polypeptide or nucleic acid of the invention in a biological sample involves obtaining a biological sample from a test subject and contacting the biological sample with a compound or an agent capable of detecting a polypeptide or nucleic acid (*e.g.*, mRNA, genomic DNA) of the invention such that the presence of a polypeptide or nucleic acid of the invention is detected in the biological sample. A preferred agent for detecting mRNA or genomic DNA encoding a polypeptide of the invention is a labeled nucleic acid probe capable of hybridizing to mRNA or genomic DNA encoding a polypeptide of the invention. The nucleic acid probe can be, for example, a full-length cDNA, such as the nucleic acid of SEQ ID NOs:1, 3, 4, 6, 7, 9, 10, 12, 13, 15, 16, 18, 19, 21, 22, 24, 25, 27, 28 or 30, or a portion thereof, such as an oligonucleotide of at least 15, 30, 50, 100, 250 or 500 nucleotides in length and sufficient to specifically hybridize under stringent conditions to a mRNA or genomic DNA encoding a polypeptide of the invention. Other suitable probes for use in the diagnostic assays of the invention are described herein.

A preferred agent for detecting a polypeptide of the invention is an antibody capable of binding to a polypeptide of the invention, preferably an antibody with a detectable label. Antibodies can be polyclonal, or more preferably, monoclonal. An intact antibody, or a fragment thereof (*e.g.*, Fab or $F(ab')_2$) can be used. The term "labeled", with regard to the probe or antibody, is intended to encompass direct labeling of the probe or antibody by coupling (i.e., physically linking) a detectable substance to the probe or antibody, as well as indirect labeling of the probe or antibody by reactivity with another reagent that is directly

labeled. Examples of indirect labeling include detection of a primary antibody using a fluorescently labeled secondary antibody and end-labeling of a DNA probe with biotin such that it can be detected with fluorescently labeled streptavidin. The term "biological sample" is intended to include tissues, cells and biological fluids isolated from a subject, as well as tissues, cells and fluids present within a subject. That is, the detection method of the invention can be used to detect mRNA, protein, or genomic DNA in a biological sample *in vitro* as well as *in vivo*. For example, *in vitro* techniques for detection of mRNA include Northern hybridizations and *in situ* hybridizations. *In vitro* techniques for detection of a polypeptide of the invention include enzyme linked immunosorbent assays (ELISAs), Western blots, immunoprecipitations and immunofluorescence. *In vitro* techniques for detection of genomic DNA include Southern hybridizations. Furthermore, *in vivo* techniques for detection of a polypeptide of the invention include introducing into a subject a labeled antibody directed against the polypeptide. For example, the antibody can be labeled with a radioactive marker whose presence and location in a subject can be detected by standard imaging techniques.

In one embodiment, the biological sample contains protein molecules from the test subject. Alternatively, the biological sample can contain mRNA molecules from the test subject or genomic DNA molecules from the test subject. A preferred biological sample is a peripheral blood leukocyte sample isolated by conventional means from a subject.

In another embodiment, the methods further involve obtaining a control biological sample from a control subject, contacting the control sample with a compound or agent capable of detecting a polypeptide of the invention or mRNA or genomic DNA encoding a polypeptide of the invention, such that the presence of the polypeptide or mRNA or genomic DNA encoding the polypeptide is detected in the biological sample, and comparing the presence of the polypeptide or mRNA or genomic DNA encoding the polypeptide in the control sample with the presence of the polypeptide or mRNA or genomic DNA encoding the polypeptide in the test sample.

The invention also encompasses kits for detecting the presence of a polypeptide or nucleic acid of the invention in a biological sample (a test sample). Such kits can be used to determine if a subject is suffering from or is at increased risk of developing a disorder associated with aberrant expression of a polypeptide of the invention (*e.g.*, a proliferative disorder, *e.g.*, psoriasis or cancer). For example, the kit can comprise a labeled compound or agent capable of detecting the polypeptide or mRNA encoding the polypeptide in a biological sample and means for determining the amount of the polypeptide or mRNA in the sample (*e.g.*, an antibody which binds the polypeptide or an oligonucleotide probe which binds to DNA or mRNA encoding the polypeptide). Kits can also include instructions for observing that the tested subject is suffering from or is at risk of developing

a disorder associated with aberrant expression of the polypeptide if the amount of the polypeptide or mRNA encoding the polypeptide is above or below a normal level.

For antibody-based kits, the kit can comprise, for example: (1) a first antibody (*e.g.*, attached to a solid support) which binds to a polypeptide of the invention; and, optionally, (2) a second, different antibody which binds to either the polypeptide or the first antibody and is conjugated to a detectable agent.

For oligonucleotide-based kits, the kit can comprise, for example: (1) an oligonucleotide, *e.g.*, a detectably labeled oligonucleotide, which hybridizes to a nucleic acid sequence encoding a polypeptide of the invention or (2) a pair of primers useful for amplifying a nucleic acid molecule encoding a polypeptide of the invention. The kit can also comprise, *e.g.*, a buffering agent, a preservative, or a protein stabilizing agent. The kit can also comprise components necessary for detecting the detectable agent (*e.g.*, an enzyme or a substrate). The kit can also contain a control sample or a series of control samples which can be assayed and compared to the test sample contained. Each component of the kit is usually enclosed within an individual container and all of the various containers are within a single package along with instructions for observing whether the tested subject is suffering from or is at risk of developing a disorder associated with aberrant expression of the polypeptide.

2. Prognostic Assays

The methods described herein can furthermore be utilized as diagnostic or prognostic assays to identify subjects having or at risk of developing a disease or disorder associated with aberrant expression or activity of a polypeptide of the invention. For example, the assays described herein, such as the preceding diagnostic assays or the following assays, can be utilized to identify a subject having or at risk of developing a disorder associated with aberrant expression or activity of a polypeptide of the invention. Alternatively, the prognostic assays can be utilized to identify a subject having or at risk for developing such a disease or disorder. Thus, the present invention provides a method in which a test sample is obtained from a subject and a polypeptide or nucleic acid (*e.g.*, mRNA, genomic DNA) of the invention is detected, wherein the presence of the polypeptide or nucleic acid is diagnostic for a subject having or at risk of developing a disease or disorder associated with aberrant expression or activity of the polypeptide. As used herein, a "test sample" refers to a biological sample obtained from a subject of interest. For example, a test sample can be a biological fluid (*e.g.*, serum), cell sample, or tissue.

Furthermore, the prognostic assays described herein can be used to determine whether a subject can be administered an agent (*e.g.*, an agonist, antagonist, peptidomimetic, protein, peptide, nucleic acid, small molecule, or other drug candidate) to

treat a disease or disorder associated with aberrant expression or activity of a polypeptide of the invention. For example, such methods can be used to determine whether a subject can be effectively treated with a specific agent or class of agents (*e.g.*, agents of a type which decrease activity of the polypeptide). Thus, the present invention provides methods for determining whether a subject can be effectively treated with an agent for a disorder associated with aberrant expression or activity of a polypeptide of the invention in which a test sample is obtained and the polypeptide or nucleic acid encoding the polypeptide is detected (*e.g.*, wherein the presence of the polypeptide or nucleic acid is diagnostic for a subject that can be administered the agent to treat a disorder associated with aberrant expression or activity of the polypeptide).

The methods of the invention can also be used to detect genetic lesions or mutations in a gene of the invention, thereby determining if a subject with the lesioned gene is at risk for a disorder characterized aberrant expression or activity of a polypeptide of the invention. In preferred embodiments, the methods include detecting, in a sample of cells from the subject, the presence or absence of a genetic lesion or mutation characterized by at least one of an alteration affecting the integrity of a gene encoding the polypeptide of the invention, or the mis-expression of the gene encoding the polypeptide of the invention. For example, such genetic lesions or mutations can be detected by ascertaining the existence of at least one of: 1) a deletion of one or more nucleotides from the gene; 2) an addition of one or more nucleotides to the gene; 3) a substitution of one or more nucleotides of the gene; 4) a chromosomal rearrangement of the gene; 5) an alteration in the level of a messenger RNA transcript of the gene; 6) an aberrant modification of the gene, such as of the methylation pattern of the genomic DNA; 7) the presence of a non-wild type splicing pattern of a messenger RNA transcript of the gene; 8) a non-wild type level of a the protein encoded by the gene; 9) an allelic loss of the gene; and 10) an inappropriate post-translational modification of the protein encoded by the gene. As described herein, there are a large number of assay techniques known in the art which can be used for detecting lesions in a gene.

In certain embodiments, detection of the lesion involves the use of a probe/primer in a polymerase chain reaction (PCR) (*see, e.g.*, U.S. Patent NOs. 4,683,195 and 4,683,202), such as anchor PCR or RACE PCR, or, alternatively, in a ligation chain reaction (LCR) (*see, e.g.*, Landegran et al., 1988, *Science* 241:1077-80; and Nakazawa et al., 1994, *Proc. Natl. Acad. Sci. USA* 91:360-4), the latter of which can be particularly useful for detecting point mutations in a gene (*see, e.g.*, Abravaya et al., 1995, *Nucleic Acids Res.* 23:675-82). This method can include the steps of collecting a sample of cells from a patient, isolating nucleic acid (*e.g.*, genomic, mRNA or both) from the cells of the sample, contacting the nucleic acid sample with one or more primers which specifically hybridize to the selected

gene under conditions such that hybridization and amplification of the gene (if present) occurs, and detecting the presence or absence of an amplification product, or detecting the size of the amplification product and comparing the length to a control sample. It is anticipated that PCR and/or LCR may be desirable to use as a preliminary amplification step in conjunction with any of the techniques used for detecting mutations described herein.

Alternative amplification methods include: self sustained sequence replication (Guatelli et al., 1990, *Proc. Natl. Acad. Sci. USA* 87:1874-78), transcriptional amplification system (Kwoh, et al., 1989, *Proc. Natl. Acad. Sci. USA* 86:1173-7), Q-Beta Replicase (Lizardi et al., 1988, *Bio/Technology* 6:1197), or any other nucleic acid amplification method, followed by the detection of the amplified molecules using techniques well known to those of skill in the art. These detection schemes are especially useful for the detection of nucleic acid molecules if such molecules are present in very low numbers.

In an alternative embodiment, mutations in a selected gene from a sample cell can be identified by alterations in restriction enzyme cleavage patterns. For example, sample and control DNA is isolated, amplified (optionally), digested with one or more restriction endonucleases, and fragment length sizes are determined by gel electrophoresis and compared. Differences in fragment length sizes between sample and control DNA indicates mutations in the sample DNA. Moreover, the use of sequence specific ribozymes (*see*, *e.g.*, U.S. Patent No. 5,498,531) can be used to score for the presence of specific mutations by development or loss of a ribozyme cleavage site.

In other embodiments, genetic mutations can be identified by hybridizing a sample and control nucleic acids, *e.g.*, DNA or RNA, to high density arrays containing hundreds or thousands of oligonucleotides probes (Cronin et al., 1996, *Human Mutation* 7:244-55; Kozal et al., 1996, *Nature Medicine* 2:753-9). For example, genetic mutations can be identified in two-dimensional arrays containing light-generated DNA probes as described in Cronin et al., *supra*. Briefly, a first hybridization array of probes can be used to scan through long stretches of DNA in a sample and control to identify base changes between the sequences by making linear arrays of sequential overlapping probes. This step allows the identification of point mutations. This step is followed by a second hybridization array that allows the characterization of specific mutations by using smaller, specialized probe arrays complementary to all variants or mutations detected. Each mutation array is composed of parallel probe sets, one complementary to the wild-type gene and the other complementary to the mutant gene.

In yet another embodiment, any of a variety of sequencing reactions known in the art can be used to directly sequence the selected gene and detect mutations by comparing the sequence of the sample nucleic acids with the corresponding wild-type (control)

sequence. Examples of sequencing reactions include those based on techniques developed by Maxim and Gilbert (1977, *Proc. Natl. Acad. Sci. USA* 74:560) or Sanger (1977, *Proc. Natl. Acad. Sci. USA* 74:5463). It is also contemplated that any of a variety of automated sequencing procedures can be utilized when performing the diagnostic assays developed by Naeve et al. (1995, *Bio/Techniques* 19:448-53), including sequencing by mass spectrometry (*see, e.g.*, PCT Publication No. WO 94/16101; Cohen et al., 1996, *Adv. Chromatogr.* 36:127-62; and Griffin et al., 1993, *Appl. Biochem. Biotechnol.* 38:147-59).

Other methods for detecting mutations in a selected gene include methods in which protection from cleavage agents is used to detect mismatched bases in RNA/RNA or RNA/DNA heteroduplexes (Myers et al., 1985, *Science* 230:1242). In general, the technique of mismatch cleavage entails providing heteroduplexes formed by hybridizing (labeled) RNA or DNA containing the wild-type sequence with potentially mutant RNA or DNA obtained from a tissue sample. The double-stranded duplexes are treated with an agent which cleaves single-stranded regions of the duplex such as which will exist due to basepair mismatches between the control and sample strands. RNA/DNA duplexes can be treated with RNase to digest mismatched regions, and DNA/DNA hybrids can be treated with S1 nuclease to digest mismatched regions.

In other embodiments, either DNA/DNA or RNA/DNA duplexes can be treated with hydroxylamine or osmium tetroxide and with piperidine in order to digest mismatched regions. After digestion of the mismatched regions, the resulting material is then separated by size on denaturing polyacrylamide gels to determine the site of mutation. *See, e.g.*, Cotton et al., 1988, *Proc. Natl. Acad. Sci. USA* 85:4397; Saleeba et al., 1992, *Methods Enzymol.* 217:286-95. In a preferred embodiment, the control DNA or RNA can be labeled for detection.

In still another embodiment, the mismatch cleavage reaction employs one or more proteins that recognize mismatched base pairs in double-stranded DNA (so called DNA mismatch repair enzymes) in defined systems for detecting and mapping point mutations in cDNAs obtained from samples of cells. For example, the mutY enzyme of *E. coli* cleaves A at G/A mismatches and the thymidine DNA glycosylase from HeLa cells cleaves T at G/T mismatches (Hsu et al., 1994, *Carcinogenesis* 15:1657-62). According to an exemplary embodiment, a probe based on a selected sequence, *e.g.*, a wild-type sequence, is hybridized to a cDNA or other DNA product from a test cell(s). The duplex is treated with a DNA mismatch repair enzyme, and the cleavage products, if any, can be detected from electrophoresis protocols or the like. *See, e.g.*, U.S. Patent No. 5,459,039.

In other embodiments, alterations in electrophoretic mobility will be used to identify mutations in genes. For example, single strand conformation polymorphism (SSCP) may be used to detect differences in electrophoretic mobility between mutant and wild type

nucleic acids (Orita et al., 1989, *Proc. Natl. Acad. Sci. USA* 86:2766; *see also* Cotton, 1993, *Mutat. Res.* 285:125-44; Hayashi, 1992, *Genet. Anal. Tech. Appl.* 9:73-9). Single-stranded DNA fragments of sample and control nucleic acids will be denatured and allowed to renature. The secondary structure of single-stranded nucleic acids varies according to sequence, and the resulting alteration in electrophoretic mobility enables the detection of even a single base change. The DNA fragments may be labeled or detected with labeled probes. The sensitivity of the assay may be enhanced by using RNA (rather than DNA), in which the secondary structure is more sensitive to a change in sequence. In a preferred embodiment, the subject method utilizes heteroduplex analysis to separate double stranded heteroduplex molecules on the basis of changes in electrophoretic mobility (Keen et al., 1991, *Trends Genet.* 7:5).

In yet another embodiment, the movement of mutant or wild-type fragments in polyacrylamide gels containing a gradient of denaturant is assayed using denaturing gradient gel electrophoresis (DGGE) (Myers et al., 1985, *Nature* 313:495). When DGGE is used as the method of analysis, DNA will be modified to insure that it does not completely denature, for example by adding a `GC clamp of approximately 40 bp of high-melting GC-rich DNA by PCR. In a further embodiment, a temperature gradient is used in place of a denaturing gradient to identify differences in the mobility of control and sample DNA (Rosenbaum and Reissner, 1987, *Biophys. Chem.* 265:12753).

Examples of other techniques for detecting point mutations include, but are not limited to, selective oligonucleotide hybridization, selective amplification, or selective primer extension. For example, oligonucleotide primers may be prepared in which the known mutation is placed centrally and then hybridized to target DNA under conditions which permit hybridization only if a perfect match is found (Saiki et al., 1986, *Nature* 324:163; Saiki et al., 1989, *Proc. Natl. Acad. Sci. USA* 86:6230). Such allele specific oligonucleotides are hybridized to PCR amplified target DNA or a number of different mutations when the oligonucleotides are attached to the hybridizing membrane and hybridized with labeled target DNA.

Alternatively, allele specific amplification technology which depends on selective PCR amplification may be used in conjunction with the instant invention. Oligonucleotides used as primers for specific amplification may carry the mutation of interest in the center of the molecule (so that amplification depends on differential hybridization; Gibbs et al., 1989, *Nucleic Acids Res.* 17:2437-48) or at the extreme 3' end of one primer where, under appropriate conditions, mismatch can prevent or reduce polymerase extension (Prossner, 1993, *Tibtech* 11:238). In addition, it may be desirable to introduce a novel restriction site in the region of the mutation to create cleavage-based detection (Gasparini et al., 1992, *Mol. Cell Probes* 6:1). It is anticipated that in certain embodiments amplification may also be

performed using Taq ligase for amplification (Barany, 1991, *Proc. Natl. Acad. Sci. USA* 88:189). In such cases, ligation will occur only if there is a perfect match at the 3' end of the 5' sequence making it possible to detect the presence of a known mutation at a specific site by looking for the presence or absence of amplification.

The methods described herein may be performed, for example, by utilizing pre-packaged diagnostic kits comprising at least one probe nucleic acid or antibody reagent described herein, which may be conveniently used, *e.g.*, in clinical settings to diagnose patients exhibiting symptoms or family history of a disease or illness involving a gene encoding a polypeptide of the invention. Furthermore, any cell type or tissue, preferably peripheral blood leukocytes, in which the polypeptide of the invention is expressed may be utilized in the prognostic assays described herein.

3. Pharmacogenomics

Agents, or modulators which have a stimulatory or inhibitory effect on activity or expression of a polypeptide of the invention as identified by a screening assay described herein can be administered to individuals to treat (prophylactically or therapeutically) disorders associated with aberrant activity of the polypeptide. In conjunction with such treatment, the pharmacogenomics (i.e., the study of the relationship between an individual's genotype and that individual's response to a foreign compound or drug) of the individual may be considered. Differences in metabolism of therapeutics can lead to severe toxicity or therapeutic failure by altering the relation between dose and blood concentration of the pharmacologically active drug. Thus, the pharmacogenomics of the individual permits the selection of effective agents (*e.g.*, drugs) for prophylactic or therapeutic treatments based on a consideration of the individual's genotype. Such pharmacogenomics can further be used to determine appropriate dosages and therapeutic regimens. Accordingly, the activity of a polypeptide of the invention, expression of a nucleic acid of the invention, or mutation content of a gene of the invention in an individual can be determined to thereby select appropriate agent(s) for therapeutic or prophylactic treatment of the individual.

Pharmacogenomics deals with clinically significant hereditary variations in the response to drugs due to altered drug disposition and abnormal action in affected persons. *See, e.g.*, Linder, 1997, *Clin. Chem.* 43(2):254-66. In general, two types of pharmacogenetic conditions can be differentiated. Genetic conditions transmitted as a single factor altering the way drugs act on the body are referred to as "altered drug action." Genetic conditions transmitted as single factors altering the way the body acts on drugs are referred to as "altered drug metabolism". These pharmacogenetic conditions can occur either as rare defects or as polymorphisms. For example, glucose-6-phosphate dehydrogenase deficiency (G6PD) is a common inherited enzymopathy in which the main

clinical complication is haemolysis after ingestion of oxidant drugs (anti-malarials, sulfonamides, analgesics, nitrofurans) and consumption of fava beans.

As an illustrative embodiment, the activity of drug metabolizing enzymes is a major determinant of both the intensity and duration of drug action. The discovery of genetic polymorphisms of drug metabolizing enzymes (*e.g.*, N-acetyltransferase 2 (NAT 2) and cytochrome P450 enzymes CYP2D6 and CYP2C19) has provided an explanation as to why some patients do not obtain the expected drug effects or show exaggerated drug response and serious toxicity after taking the standard and safe dose of a drug. These polymorphisms are expressed in two phenotypes in the population, the extensive metabolizer (EM) and poor metabolizer (PM). The prevalence of PM is different among different populations. For example, the gene coding for CYP2D6 is highly polymorphic and several mutations have been identified in PM, which all lead to the absence of functional CYP2D6. Poor metabolizers of CYP2D6 and CYP2C19 quite frequently experience exaggerated drug response and side effects when they receive standard doses. If a metabolite is the active therapeutic moiety, a PM will show no therapeutic response, as demonstrated for the analgesic effect of codeine mediated by its CYP2D6-formed metabolite morphine. The other extreme are the so called ultra-rapid metabolizers who do not respond to standard doses. Recently, the molecular basis of ultra-rapid metabolism has been identified to be due to CYP2D6 gene amplification.

Thus, the activity of a polypeptide of the invention, expression of a nucleic acid encoding the polypeptide, or mutation content of a gene encoding the polypeptide in an individual can be determined to thereby select appropriate agent(s) for therapeutic or prophylactic treatment of the individual. In addition, pharmacogenetic studies can be used to apply genotyping of polymorphic alleles encoding drug-metabolizing enzymes to the identification of an individual's drug responsiveness phenotype. This knowledge, when applied to dosing or drug selection, can avoid adverse reactions or therapeutic failure and thus enhance therapeutic or prophylactic efficiency when treating a subject with a modulator of activity or expression of the polypeptide, such as a modulator identified by one of the exemplary screening assays described herein.

4. Monitoring of Effects During Clinical Trials

Monitoring the influence of agents (*e.g.*, drugs, compounds) on the expression or activity of a polypeptide of the invention (*e.g.*, the ability to modulate aberrant cell proliferation chemotaxis, and/or differentiation) can be applied not only in basic drug screening, but also in clinical trials. For example, the effectiveness of an agent, as determined by a screening assay as described herein, to increase gene expression, protein levels or protein activity, can be monitored in clinical trials of subjects exhibiting decreased

gene expression, protein levels, or protein activity. Alternatively, the effectiveness of an agent, as determined by a screening assay, to decrease gene expression, protein levels or protein activity, can be monitored in clinical trials of subjects exhibiting increased gene expression, protein levels, or protein activity. In such clinical trials, expression or activity of a polypeptide of the invention and preferably, that of other polypeptide that have been implicated in for example, a cellular proliferation disorder, can be used as a marker of the immune responsiveness of a particular cell.

For example, and not by way of limitation, genes, including those of the invention, that are modulated in cells by treatment with an agent (*e.g.*, compound, drug or small molecule) which modulates activity or expression of a polypeptide of the invention (*e.g.*, as identified in a screening assay described herein) can be identified. Thus, to study the effect of agents on cellular proliferation disorders, for example, in a clinical trial, cells can be isolated and RNA prepared and analyzed for the levels of expression of a gene of the invention and other genes implicated in the disorder. The levels of gene expression (i.e., a gene expression pattern) can be quantified by Northern blot analysis or RT-PCR, as described herein, or alternatively by measuring the amount of protein produced, by one of the methods as described herein, or by measuring the levels of activity of a gene of the invention or other genes. In this way, the gene expression pattern can serve as a marker, indicative of the physiological response of the cells to the agent. Accordingly, this response state may be determined before, and at various points during, treatment of the individual with the agent.

In a preferred embodiment, the present invention provides a method for monitoring the effectiveness of treatment of a subject with an agent (*e.g.*, an agonist, antagonist, peptidomimetic, protein, peptide, nucleic acid, small molecule, or other drug candidate identified by the screening assays described herein) comprising the steps of (i) obtaining a pre-administration sample from a subject prior to administration of the agent; (ii) detecting the level of the polypeptide or nucleic acid of the invention in the preadministration sample; (iii) obtaining one or more post-administration samples from the subject; (iv) detecting the level the of the polypeptide or nucleic acid of the invention in the post-administration samples; (v) comparing the level of the polypeptide or nucleic acid of the invention in the pre-administration sample with the level of the polypeptide or nucleic acid of the invention in the post-administration sample or samples; and (vi) altering the administration of the agent to the subject accordingly. For example, increased administration of the agent may be desirable to increase the expression or activity of the polypeptide to higher levels than detected, i.e., to increase the effectiveness of the agent. Alternatively, decreased administration of the agent may be desirable to decrease expression or activity of the polypeptide to lower levels than detected, i.e., to decrease the effectiveness of the agent.

C. Methods of Treatment

The present invention provides for both prophylactic and therapeutic methods of treating a subject at risk of (or susceptible to) a disorder or having a disorder associated with aberrant expression or activity of a polypeptide of the invention, *e.g.*, cardiac infection (*e.g.*, myocarditis or dilated cardiomyopathy), central nervous system infection (*e.g.*, non-specific febrile illness or meningoencephalitis), pancreatic infection (*e.g.*, acute pancreatitis), respiratory infection (pneumonia), gastrointestinal infection, type I diabetes, cancer, familia hypercholesterolemia, treat hemophilia B, Marfan syndrome, protein S deficiency, allergy, inflammation, and gastroduodenal ulcer. Moreover, the polypeptides of the invention can be used to modulate cellular function, survival, morphology, proliferation and/or differentiation.

1. Prophylactic Methods

In one aspect, the invention provides a method for preventing in a subject, a disease or condition associated with an aberrant expression or activity of a polypeptide of the invention, by administering to the subject an agent which modulates expression or at least one activity of the polypeptide. Subjects at risk for a disease which is caused or contributed to by aberrant expression or activity of a polypeptide of the invention can be identified by, for example, any or a combination of diagnostic or prognostic assays as described herein. Administration of a prophylactic agent can occur prior to the manifestation of symptoms characteristic of the aberrancy, such that a disease or disorder is prevented or, alternatively, delayed in its progression. Depending on the type of aberrancy, for example, an agonist or antagonist agent can be used for treating the subject.

2. Therapeutic Methods

Another aspect of the invention pertains to methods of modulating expression or activity of a polypeptide of the invention for therapeutic purposes. The modulatory method of the invention involves contacting a cell with an agent that modulates one or more of the activities of the polypeptide. An agent that modulates activity can be an agent as described herein, such as a nucleic acid or a protein, a naturally-occurring cognate ligand of the polypeptide, a peptide, a peptidomimetic, or other small molecule. In one embodiment, the agent stimulates one or more of the biological activities of the polypeptide. Examples of such stimulatory agents include the active polypeptide of the invention and a nucleic acid molecule encoding the polypeptide of the invention that has been introduced into the cell. In another embodiment, the agent inhibits one or more of the biological activities of the polypeptide of the invention. Examples of such inhibitory agents include antisense nucleic acid molecules and antibodies. These modulatory methods can be performed *in vitro* (*e.g.*,

by culturing the cell with the agent) or, alternatively, *in vivo* (*e.g.*, by administering the agent to a subject). As such, the present invention provides methods of treating an individual afflicted with a disease or disorder characterized by aberrant expression or activity of a polypeptide of the invention. In one embodiment, the method involves administering an agent (*e.g.*, an agent identified by a screening assay described herein), or combination of agents that modulates (*e.g.*, upregulates or downregulates) expression or activity. In another embodiment, the method involves administering a polypeptide of the invention or a nucleic acid molecule of the invention as therapy to compensate for reduced or aberrant expression or activity of the polypeptide.

Stimulation of activity is desirable in situations in which activity or expression is abnormally low or downregulated and/or in which increased activity is likely to have a beneficial effect. Conversely, inhibition of activity is desirable in situations in which activity or expression is abnormally high or upregulated and/or in which decreased activity is likely to have a beneficial effect.

The contents of all references, patents and published patent applications cited throughout this application are hereby incorporated by reference.

Deposit of Clones

Clones containing cDNA molecules encoding human MANGO 003 were deposited with the American Type Culture Collection (ATCC® 10801 University Boulevard, Manassas, VA 20110-2209) on March 30, 1999 as Accession Number 207178, as part of a composite deposit representing a mixture of three strains, each carrying one recombinant plasmid harboring a particular cDNA clone.

To distinguish the strains and isolate a strain harboring a particular cDNA clone, an aliquot of the mixture can be streaked out to single colonies on nutrient medium (*e.g.*, LB plates) supplemented with 100 g/ml ampicillin, single colonies grown, and then plasmid DNA extracted using a standard minipreparation procedure. Next, a sample of the DNA minipreparation can be digested with a combination of the restriction enzymes *Sal* I and *Not* I, and the resultant products resolved on a 0.8% agarose gel using standard DNA electrophoresis conditions. The digest liberates fragments as follows:

human MANGO 003 (clone EpthLa6a1): 3.2 kB

The identity of the strains can be inferred from the fragments liberated.

Clones containing cDNA molecules encoding human INTERCEPT 340, MANGO 347, and TANGO 272 were deposited with the American Type Culture Collection (ATCC®

10801 University Boulevard, Manassas, VA 20110-2209) on June 18, 1999 as Accession Number PTA-250, as part of a composite deposit representing a mixture of three strains, each carrying one recombinant plasmid harboring a particular cDNA clone.

To distinguish the strains and isolate a strain harboring a particular cDNA clone, an aliquot of the mixture can be streaked out to single colonies on nutrient medium (*e.g.*, LB plates) supplemented with 100 g/ml ampicillin, single colonies grown, and then plasmid DNA extracted using a standard minipreparation procedure. Next, a sample of the DNA minipreparation can be digested with a combination of the restriction enzymes *Sal* I and *Not* I, and the resultant products resolved on a 0.8% agarose gel using standard DNA electrophoresis conditions. The digest liberates fragments as follows:

human INTERCEPT 340 (clone EpI340): 3.3 kB
human MANGO 347 (clone EpM347): 1.4 kB
human TANGO 272 (clone EpT272): 5.0 kB

The identity of the strains can be inferred from the fragments liberated.

Clones containing cDNA molecules encoding human TANGO 295, TANGO 354, and TANGO 378 were deposited with the American Type Culture Collection (ATCC® 10801 University Boulevard, Manassas, VA 20110-2209) on June 18, 1999 as Accession Number PTA-249, as part of a composite deposit representing a mixture of three strains, each carrying one recombinant plasmid harboring a particular cDNA clone.

To distinguish the strains and isolate a strain harboring a particular cDNA clone, an aliquot of the mixture can be streaked out to single colonies on nutrient medium (*e.g.*, LB plates) supplemented with 100 g/ml ampicillin, single colonies grown, and then plasmid DNA extracted using a standard minipreparation procedure. Next, a sample of the DNA minipreparation can be digested with a combination of the restriction enzymes *Sal* I and *Not* I, and the resultant products resolved on a 0.8% agarose gel using standard DNA electrophoresis conditions. The digest liberates fragments as follows:

human TANGO 295 (clone EpT295): 1.5 kB
human TANGO 354 (clone EpT354): 1.8 kB
human TANGO 378 (clone EpT378): 3.3 kB

The identity of the strains can be inferred from the fragments liberated.

All publications, patents and patent applications mentioned in this specification are herein incorporated by reference into the specification to the same extent as if each individual publication, patent or patent application was specifically and individually indicated to be incorporated herein by reference.

Equivalents

Those skilled in the art will recognize, or be able to ascertain using no more than routine experimentation, many equivalents to the specific embodiments of the invention described herein. Such equivalents are intended to be encompassed by the following Claims.

International Application No: PCT/ /

MICROORGANISMS

Optional Sheet in connection with the microorganism referred to on pages ___, lines _____ of the description [1]

A. IDENTIFICATION OF DEPOSIT [2]

Further deposits are identified on an additional sheet [3]

Name of depositary institution [4]

American Type Culture Collection

Address of depositary institution (including postal code and country) [4]

10801 University Blvd.
Manassas, VA 20110-2209
US

Date of deposit [5] March 30, 1999 Accession Number [6] 207178

B. ADDITIONAL INDICATIONS [7] (leave blank if not applicable). This information is continued on a separate attached sheet

C. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE [8] (if the indications are not all designated States)

D. SEPARATE FURNISHING OF INDICATIONS [9] (leave blank if not applicable)

The indications listed below will be submitted to the International Bureau later [9] (Specify the general nature of the indications e.g., "Accession Number of Deposit")

E. ☑ This sheet was received with the International application when filed (to be checked by the receiving Office)

(Authorized Officer)

☐ The date of receipt (from the applicant) by the International Bureau [10]

was

(Authorized Officer)

Form PCT/RO/134 (January 1981)

International Application No: PCT/ /

Form PCT/RO/134 (cont.)

American Type Culture Collection

10801 University Blvd.
Manassas, VA 20110-2209
US

Accession No.	Date of Deposit
PTA-249	June 18, 1999
PTA-250	June 18, 1999

What is claimed is:

1. An isolated nucleic acid molecule selected from the group consisting of:

a) a nucleic acid molecule comprising a nucleotide sequence which is at least 55% identical to the nucleotide sequence of SEQ ID NOs:1, 3, 4, 6, 7, 9, 10, 12, 13, 15, 16, 18, 19, 21, 22, 24, 25, 27, 28, 30, the cDNA insert of the plasmid deposited with the ATCC® as Accession Number 207178, the cDNA insert of the plasmid deposited with the ATCC® as Accession Number PTA-249, the cDNA insert of the plasmid deposited with the ATCC® as Accession Number PTA-250, or a complement thereof;

b) a nucleic acid molecule comprising a fragment of at least 300 nucleotides of the nucleotide sequence of SEQ ID NOs:1, 3, 4, 6, 7, 9, 10, 12, 13, 15, 16, 18, 19, 21, 22, 24, 25, 27, 28, 30, the cDNA insert of the plasmid deposited with the ATCC® as Accession Number 207178, the cDNA insert of the plasmid deposited with the ATCC® as Accession Number PTA-249, the cDNA insert of the plasmid deposited with the ATCC® as Accession Number PTA-250, or a complement thereof;

c) a nucleic acid molecule which encodes a polypeptide comprising the amino acid sequence of SEQ ID NOs:2, 5, 8, 11, 14, 17, 20, 23, 26, 29, the amino acid sequence encoded by the cDNA insert of the plasmid deposited with the ATCC® as Accession Number 207178, the amino acid sequence encoded by the cDNA insert of the plasmid deposited with the ATCC® as Accession Number PTA-249, or the amino acid sequence encoded by the cDNA insert of the plasmid deposited with the ATCC® as Accession Number PTA-250;

d) a nucleic acid molecule which encodes a fragment of a polypeptide comprising the amino acid sequence of SEQ ID NOs:2, 5, 8, 11, 14, 17, 20, 23, 26, 29, the amino acid sequence encoded by the cDNA insert of the plasmid deposited with the ATCC® as Accession Number 207178, the amino acid sequence encoded by the cDNA insert of the plasmid deposited with the ATCC® as Accession Number PTA-249, or the amino acid sequence encoded by the cDNA insert of the plasmid deposited with the ATCC® as Accession Number PTA-250, wherein the fragment comprises at least 15 contiguous amino acids of SEQ ID NOs:2, 5, 8, 11, 14, 17, 20, 23, 26, 29, the amino acid sequence encoded by the cDNA insert of the plasmid deposited with the ATCC® as Accession Number 207178, the amino acid sequence encoded by the cDNA insert of the plasmid deposited with the ATCC® as Accession Number PTA-249, or the amino acid sequence encoded by the cDNA insert of the plasmid deposited with the ATCC® as Accession Number PTA-250; and

e) a nucleic acid molecule which encodes a naturally occurring allelic variant of a polypeptide comprising the amino acid sequence of SEQ ID NOs:2, 5, 8, 11, 14, 17, 20,

23, 26, 29, the amino acid sequence encoded by the cDNA insert of the plasmid deposited with the ATCC® as Accession Number 207178, the amino acid sequence encoded by the cDNA insert of the plasmid deposited with the ATCC® as Accession Number PTA-249, or the amino acid sequence encoded by the cDNA insert of the plasmid deposited with the ATCC® as Accession Number PTA-250, wherein the nucleic acid molecule hybridizes to a nucleic acid molecule comprising SEQ ID NOs:1, 3, 4, 6, 7, 9, 10, 12, 13, 15, 16, 18, 19, 21, 22, 24, 25, 27, 28, 30, or a complement thereof, under stringent conditions.

2. The isolated nucleic acid molecule of Claim 1, which is selected from the group consisting of:

a) a nucleic acid comprising the nucleotide sequence of SEQ ID NOs:1, 3, 4, 6, 7, 9, 10, 12, 13, 15, 16, 18, 19, 21, 22, 24, 25, 27, 28, 30, the cDNA insert of the plasmid deposited with the ATCC® as Accession Number 207178, the cDNA insert of the plasmid deposited with the ATCC® as Accession Number PTA-249, the cDNA insert of the plasmid deposited with the ATCC® as Accession Number PTA-250, or a complement thereof; and

b) a nucleic acid molecule which encodes a polypeptide comprising the amino acid sequence of SEQ ID NOs:2, 5, 8, 11, 14, 17, 20, 23, 26, 29, the amino acid sequence encoded by the cDNA insert of the plasmid deposited with the ATCC® as Accession Number 207178, the amino acid sequence encoded by the cDNA insert of the plasmid deposited with the ATCC® as Accession Number PTA-249, or the amino acid sequence encoded by the cDNA insert of the plasmid deposited with the ATCC® as Accession Number PTA-250.

3. The nucleic acid molecule of Claim 1 further comprising vector nucleic acid sequences.

4. The nucleic acid molecule of Claim 1 further comprising nucleic acid sequences encoding a heterologous polypeptide.

5. A host cell which contains the nucleic acid molecule of Claim 1.

6. The host cell of Claim 5 which is a mammalian host cell.

7. A non-human mammalian host cell containing the nucleic acid molecule of Claim 1.

8. An isolated polypeptide selected from the group consisting of:

a) a fragment of a polypeptide comprising the amino acid sequence of SEQ ID NOs:2, 5, 8, 11, 14, 17, 20, 23, 26, or 29, wherein the fragment comprises at least 15 contiguous amino acids of SEQ ID NOs:2, 5, 8, 11, 14, 17, 20, 23, 26, or 29;

b) a naturally occurring allelic variant of a polypeptide comprising the amino acid sequence of SEQ ID NOs:2, 5, 8, 11, 14, 17, 20, 23, 26, or 29, the amino acid sequence encoded by the cDNA insert of the plasmid deposited with the ATCC® as Accession Number 207178, the amino acid sequence encoded by the cDNA insert of the plasmid deposited with the ATCC® as Accession Number PTA-249, or the amino acid sequence encoded by the cDNA insert of the plasmid deposited with the ATCC® as Accession Number PTA-250, wherein the polypeptide is encoded by a nucleic acid molecule which hybridizes to a nucleic acid molecule comprising SEQ ID NOs: 1, 3, 4, 6, 7, 9, 10, 12, 13, 15, 16, 18, or a complement thereof under stringent conditions; and

c) a polypeptide which is encoded by a nucleic acid molecule comprising a nucleotide sequence which is at least 55% identical to a nucleic acid comprising the nucleotide sequence of SEQ ID NOs:1, 3, 4, 6, 7, 9, 10, 12, 13, 15, 16, 18, 19, 21, 22, 24, 25, 27, 28, 30, or a complement thereof.

9. The isolated polypeptide of Claim 8 comprising the amino acid sequence of SEQ ID NOs:2, 5, 8, 11, 14, 17, 20, 23, 26, or 29.

10. The polypeptide of Claim 8 further comprising heterologous amino acid sequences.

11. An antibody which selectively binds to a polypeptide of Claim 8.

12. A method for producing a polypeptide selected from the group consisting of:

a) a polypeptide comprising the amino acid sequence of SEQ ID NOs:2, 5, 8, 11, 14, 17, 20, 23, 26, or 29, the amino acid sequence encoded by the cDNA insert of the plasmid deposited with the ATCC® as Accession Number 207178, the amino acid sequence encoded by the cDNA insert of the plasmid deposited with the ATCC® as Accession Number PTA-249, or the amino acid sequence encoded by the cDNA insert of the plasmid deposited with the ATCC® as Accession Number PTA-250;

b) a polypeptide comprising a fragment of the amino acid sequence of SEQ ID NOs:2, 5, 8, 11, 14, 17, 20, 23, 26, or 29, the amino acid sequence encoded by the cDNA insert of the plasmid deposited with the ATCC® as Accession Number 207178, the amino acid sequence encoded by the cDNA insert of the plasmid deposited with the ATCC® as Accession Number PTA-249, or the amino acid sequence encoded by the cDNA insert of the plasmid deposited with the ATCC® as Accession Number PTA-250, wherein the

fragment comprises at least 15 contiguous amino acids of SEQ ID NOs:2, 5, 8, 11, 14, 17, 20, 23, 26, or 29, the amino acid sequence encoded by the cDNA insert of the plasmid deposited with the ATCC® as Accession Number 207178, the amino acid sequence encoded by the cDNA insert of the plasmid deposited with the ATCC® as Accession Number PTA-249, or the amino acid sequence encoded by the cDNA insert of the plasmid deposited with the ATCC® as Accession Number PTA-250; and

c) a naturally occurring allelic variant of a polypeptide comprising the amino acid sequence of SEQ ID NOs:2, 5, 8, 11, 14, 17, 20, 23, 26, or 29, the amino acid sequence encoded by the cDNA insert of the plasmid deposited with the ATCC® as Accession Number 207178, the amino acid sequence encoded by the cDNA insert of the plasmid deposited with the ATCC® as Accession Number PTA-249, or the amino acid sequence encoded by the cDNA insert of the plasmid deposited with the ATCC® as Accession Number PTA-250, wherein the polypeptide is encoded by a nucleic acid molecule which hybridizes to a nucleic acid molecule comprising SEQ ID NOs:1, 3, 4, 6, 7, 9, 10, 12, 13, 15, 16, 18, 19, 21, 22, 24, 25, 27, 28, 30, or a complement thereof under stringent conditions;

comprising culturing the host cell of Claim 5 under conditions in which the nucleic acid molecule is expressed.

13. A method for detecting the presence of a polypeptide of Claim 8 in a sample, comprising:

a) contacting the sample with a compound which selectively binds to a polypeptide of Claim 8; and

b) determining whether the compound binds to the polypeptide in the sample.

14. The method of Claim 13, wherein the compound which binds to the polypeptide is an antibody.

15. A kit comprising a compound which selectively binds to a polypeptide of Claim 8 and instructions for use.

16. A method for detecting the presence of a nucleic acid molecule of Claim 1 in a sample, comprising the steps of:

a) contacting the sample with a nucleic acid probe or primer which selectively hybridizes to the nucleic acid molecule; and

b) determining whether the nucleic acid probe or primer binds to a nucleic acid molecule in the sample.

17. The method of Claim 16, wherein the sample comprises mRNA molecules and is contacted with a nucleic acid probe.

18. A kit comprising a compound which selectively hybridizes to a nucleic acid molecule of Claim 1 and instructions for use.

19. A method for identifying a compound which binds to a polypeptide of Claim 8 comprising the steps of:

a) contacting a polypeptide, or a cell expressing a polypeptide of Claim 8 with a test compound; and

b) determining whether the polypeptide binds to the test compound.

20. The method of Claim 19, wherein the binding of the test compound to the polypeptide is detected by a method selected from the group consisting of:

a) detection of binding by direct detecting of test compound/polypeptide binding;

b) detection of binding using a competition binding assay;

c) detection of binding using an assay for INTERCEPT 340-, MANGO 003-, MANGO 347-, TANGO 272-, TANGO 295-, TANGO 354-, or TANGO 378-mediated signal transduction.

21. A method for modulating the activity of a polypeptide of Claim 8 comprising contacting a polypeptide or a cell expressing a polypeptide of Claim 8 with a compound which binds to the polypeptide in a sufficient concentration to modulate the activity of the polypeptide.

22. A method for identifying a compound which modulates the activity of a polypeptide of Claim 8, comprising:

a) contacting a polypeptide of Claim 8 with a test compound; and

b) determining the effect of the test compound on the activity of the polypeptide to thereby identify a compound which modulates the activity of the polypeptide.

Input file I340Athsa102b12; Output File I340Athsa102b12.pat
Sequence length 3284

```
GTCGACCCACGCGTCCGTTATGTAACTATACATTTTCCCAGAAATTTTAGTATATGATATGATTTTGTTTTCTTTCATC  79

CCTTTTCCCAAGCAGTTTATTATGAAAATTTTCAAACATACAGCAATGTTGAGAAAATTTTACAGTAAATGCCTATACC 158

CATTACCTAAATTTTACCATTAACATTTTACCCTGCTGGCATTATTGTGCTTATCCATCTACGTATCCCTCTCTCCCTT 237

CATTGGTGTATTTCTAAGTAAATTGTAGGCCTCAGTACACTTCCTTCTGAATTCTTCAGCATGCACAACAGTATTATAT 316

TCCATTTTTAAAAGAGCAATTCTTGATAGATTTATATAGTTTTGTAAAATGTTCATATAGAGCTACAAATTTTATCTTT 395

TTGTTTCTTATTGTATGTCTAGGGTCCTGAAGGGGATGCTGGCATTGTTGGGATATCAGGTCCTAAAGGTCCTATTGGA 474

CACAGAGGAAACACTGGTCCCCTTGGCAGAGAAGGTATAATAGGCCCAACAGGTAGAACTGGACCCAGAGGTGAAAAGG 553

GCTTTAGAGGTGAAACTGGTCCTCAAGGACCAAGAGGTCAACCAGGGCCTCCAGGTCCACCTGGAGCACCAGGCCCAAG 632

AAAGCAAATGGATATCAATGCTGCTATTCAAGCCTTGATTGAATCAAATACTGCCCTACAGATGGAGGTAACATATCTG 711

GTTTTATTTATATTGGCACTGTCTCTCAATATACCAATTAAACAGAGAAAATTTTTGGAGGCCAAAATGTGACATTATC 790

TCAAAGATTGTATTTAAAACAGATTGAAAATGTGAAACCATTCTCAAGAACAAAGTAAGTGATTTTGGTATAATTAAAC 869

AGAAATATATGCGTAGGATGTTTTGTAAGGAAAACATTTAAATCAAAAATTTAGTACTGTTATTTGTAAGGAATTTGGT 948

ACTATCCAAGAAAGTAGTTAAATGAGGTTAGCCATGTTTCTTAAAATGAGATATATATATTATCACTACTCATTTATTT 1027

AAACTCTAATGATTCAATGTGTAATTTAAAAAACATAATACAGTAGACATAGCAATTCTTATGTTAGCTTGAAAACTAA 1106

ACTTGCAAATGTGAATTTAACCTCTTTAAAAGATTAAGGTTATTAAAGCATACACATATGCCTATGCTTAAATATAAAC 1185

                                        M  E  T  H  S  S  P  A  L  A      10
TGTTCTTTACATTCTACTCACAACTTACTACACATA ATG GAA ACA CAT TCT TCT CCT GCC TTG GCC  1251

 H  V  G  P  Q  D  F  F  V  Y  I  I  L  M  M  T  W  Q  S  Y      30
CAT GTT GGT CCT CAG GAT TTT TTT GTT TAT ATA ATT CTT ATG ATG ACT TGG CAG AGC TAC 1311

 Q  N  T  E  V  T  L  I  D  H  S  E  E  I  F  K  T  L  N  Y      50
CAG AAT ACT GAA GTG ACT TTA ATT GAC CAC AGT GAA GAG ATA TTC AAA ACC CTG AAC TAC 1371
```

FIG.1A

```
 L  S  N  L  L  H  S  I  K  N  P  L  G  T  R  D  N  P  A  R   70
CTT AGC AAT TTA TTG CAC AGC ATC AAG AAT CCT CTT GGC ACA CGA GAT AAC CCA GCA CGA  1431

 I  C  K  D  L  L  N  C  E  Q  K  V  S  D  G  K  Y  W  I  D   90
ATC TGC AAA GAT TTA CTT AAC TGT GAA CAA AAA GTA TCA GAT GGA AAA TAC TGG ATT GAC  1491

 P  N  L  G  C  P  S  D  A  I  E  V  F  C  N  F  S  A  G  G  110
CCA AAT CTT GGC TGT CCT TCA GAT GCC ATT GAG GTT TTC TGC AAT TTC AGT GCT GGT GGC  1551

 Q  T  C  L  P  P  V  S  V  T  K  L  E  F  G  V  G  K  V  Q  130
CAG ACA TGC TTA CCT CCT GTT TCT GTA ACA AAG TTG GAG TTT GGA GTT GGG AAA GTC CAG  1611

 M  N  F  L  H  L  L  S  S  E  A  T  H  I  I  T  I  H  C  L  150
ATG AAC TTC CTT CAT TTA CTG AGT TCG GAA GCC ACC CAT ATC ATC ACC ATT CAC TGT CTA  1671

 N  T  P  R  W  T  S  T  Q  T  S  G  P  G  L  P  I  G  F  K  170
AAC ACC CCA AGG TGG ACA AGC ACA CAA ACA AGT GGC CCA GGA TTG CCT ATT GGT TTC AAG  1731

 G  W  N  G  Q  I  F  K  V  N  T  L  L  E  P  K  V  L  S  D  190
GGA TGG AAT GGC CAG ATT TTT AAA GTA AAC ACT CTA CTT GAA CCT AAA GTG CTT TCA GAT  1791

 D  C  K  I  Q  D  G  S  W  H  K  A  T  F  L  F  H  T  Q  E  210
GAC TGC AAG ATT CAA GAT GGC AGC TGG CAT AAG GCA ACA TTT CTT TTT CAC ACC CAG GAA  1851

 P  N  Q  L  P  V  I  E  V  Q  K  L  P  H  L  K  T  E  R  K  230
CCT AAT CAA CTT CCA GTG ATT GAA GTA CAA AAA CTT CCT CAT CTC AAA ACT GAA CGA AAG  1911

 Y  Y  I  D  S  S  S  V  C  F  L  *                          242
TAT TAC ATT GAC AGC AGT TCT GTA TGC TTT CTG TAA                1947

AGTCTCTGAATTAGTTCCGAATTCAGGCTGTTGGCCAGGTAATTGCTGCAGAGGGAGAAATAAGACAGACAGATACAGT  2026

CATTATGAAATGCATGTAATAAAGCATTGGCTAAATCTTAAAGAATCTCAGGAAGAACAGACTTCCTCCTAAGAAGGAG  2105

AAAAGGCATTTTTAAAGGACTATGATTGATAAAGTATTTAATTCTTTTAAAAATTATATTCATCTCAGCTTTCTTAGAG  2184

AATTCCCTAGAACTAAAAATTTATAAATATGGAATTCTTCAGGGTATCTTATATTTTTGACTGAGTGCGTAGTACCCAT  2263

TAGACAGCTGGAGATGCAGAGCACTATGGAGCAATACTGGCTAATGCTTCCAGATGTGCACTGCTTCTGTCTAAAAATT  2342

ACAAGCCACAGTCTAATATGTCTTATTTTCCAAAACACTAAGCTGTATTCAGGTCCCCGATGGGCATATACATCTTAGC  2421
```

FIG.1B

CGGTGATACACTACCTCTTACGTGTTGCCTCTTTGTGTTGCTTGGTGCTCTTTCGAAAACAAGGTGCTTATGGCTTTCA 2500

TAGACTATTTCCTTTTTCATCTTTGTCATTCTTTAAAAGTGTATGTACTGGTTACATCAAGATATGTTTTGGTTGTTAG 2579

TACTTATTTTAATTTGTTTGGTCACACACTTAATAACACATGAAACTATTTATGTGAAGTCCTTGTTTTATTTTAAAAT 2658

TCTCTTTGTGTATTTGGAATCAAAGCCAGCACATTGTAACCTGTGCTTGTACGCAAAAGAATTAGATTTCTTTGTTTTT 2737

GTTTTATTTTTTAAATTGTTGTAAAAATTATTATAGGCCAGCTACATCTAGTAGTAGGTTTGGGGTACAGATTGGGGGT 2816

TGTGCCATACTGTTTTTAAAGTTCATGATCATCTGGAATGATACTTAGTGTATATATATTTTGTAAAGTTTTAATTCAG 2895

CAAATTTTTTGAAATTGCTGCTGTTTTAAATTATAAAACCTTTATATTTCTGCTTTGTAGAAATTATATGTTTTGTAGT 2974

ATTCATTGATTTTCTTTCACTGTACTTAAATTTAGTGTTAGTACTTTAAAATTTTTAATTTACCAGTCTTTAAAGCAAC 3053

ATCCAGAAAAAAAAAAAGTCTTTTCCCATTTAAAATAGGCTCAGCCAGTTCAATGTCGCCTTGTTATCAGAGAAATATTA 3132

GTTCAATACTGAAAGAAAAATATTATACCTCTTGGTATCTAGAAAAGCTTGTTCATCCATTATAAATATATCTTTAGCC 3211

ACAGCAAACCACACTTAACCTATCTATAATAAAAATGTGCTTTAAATAAAAAAAAAAAAAAAAAGGGCGGCCG 3284

FIG.1C

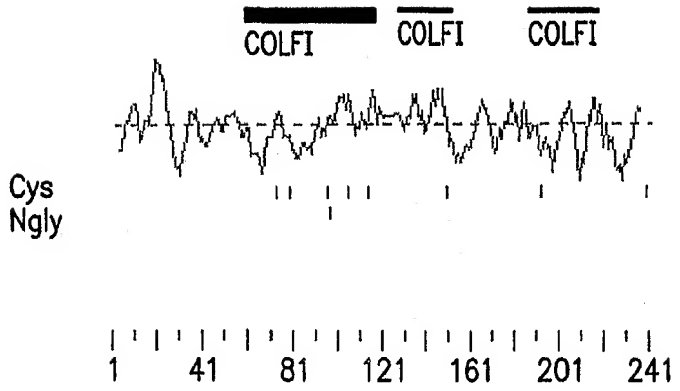


FIG.2

```
COLFI: domain 1 of 3, from 58 to 116: score 110.3, E = 1.3e-42

                    *->lksPeGksrknPARtCkDLfLchpefksGeYWiDPNqGCikDAikVf
                       +k+P+G +r+nPAR CkDL   c +  ++G YWiDPN+GC+ DAi+Vf
      INT340     58    IKNPLG-TRDNPARICKDLLNCEQKVSDGKYWIDPNLGCPSDAIEVF 103

                    CnkrfetGvgeTCisp<-*
                    Cn   f +G g+TC +p
      INT340    104 CN--FSAG-GQTCLPP    116


COLFI: domain 2 of 3, from 126 to 151: score 9.7, E = 0.0028

                    *->isnvQlTFLRLLSteAsQNiTYhCKN<-*
                       +++vQ+ FL LLS+eA   iT hC N
      INT340    126    VGKVQMNFLHLLSSEATHIITIHCLN    151


COLFI: domain 3 of 3, from 186 to 217: score 5.8, E = 0.09

                    *->tvlGeDGCssrlgewgKTViEyeTkKtlRLPIv<-*
                       +vl  D C+   g w K+ + + T+ + +LP +
      INT340    186    KVL-SDDCKIQDGSWHKATFLFHTQEPNQLPVI    217
```

FIG.3

Input file M003Athya30d3; Output File M003AthYa30d3.pat
Sequence length 3169

```
                                                                M  T  P  S  P      5
GTCGACCCACGCGTCCGCGCCCGCTGAGCCCCCCGCCGAGGTCCGGACAGGCCGAG ATG ACG CCG AGC CCC     71

 L  L  L  L  L  L  P  P  L  L  L  G  A  F  P  P  A  A  A  A     25
CTG TTG CTG CTC CTG CTG CCG CCG CTG CTG CTG GGG GCC TTC CCG CCG GCC GCC GCC GCC   131

 R  G  P  P  K  M  A  D  K  V  V  P  R  Q  V  A  R  L  G  R     45
CGA GGC CCC CCA AAG ATG GCG GAC AAG GTG GTC CCA CGG CAG GTG GCC CGG CTG GGC CGC   191

 T  V  R  L  Q  C  P  V  E  G  D  P  P  P  L  T  M  W  T  K     65
ACT GTG CGG CTG CAG TGC CCA GTG GAG GGG GAC CCG CCG CCG CTG ACC ATG TGG ACC AAG   251

 D  G  R  T  I  H  S  G  W  S  R  F  R  V  L  P  Q  G  L  K     85
GAT GGC CGC ACC ATC CAC AGC GGC TGG AGC CGC TTC CGC GTG CTG CCG CAG GGG CTG AAG   311

 V  K  Q  V  E  R  E  D  A  G  V  Y  V  C  K  A  T  N  G  F    105
GTG AAG CAG GTG GAG CGG GAG GAT GCC GGC GTG TAC GTG TGC AAG GCC ACC AAC GGC TTC   371

 G  S  L  S  V  N  Y  T  L  V  V  L  D  D  I  S  P  G  K  E    125
GGC AGC CTG AGC GTC AAC TAC ACC CTC GTC GTG CTG GAT GAC ATT AGC CCA GGG AAG GAG   431

 S  L  G  P  D  S  S  S  G  G  Q  E  D  P  A  S  Q  Q  W  A    145
AGC CTG GGG CCC GAC AGC TCC TCT GGG GGT CAA GAG GAC CCC GCC AGC CAG CAG TGG GCA   491

 R  P  R  F  T  Q  P  S  K  M  R  R  R  V  I  A  R  P  V  G    165
CGA CCG CGC TTC ACA CAG CCC TCC AAG ATG AGG CGC CGG GTG ATC GCA CGG CCC GTG GGT   551

 S  S  V  R  L  K  C  V  A  S  G  H  P  R  P  D  I  T  W  M    185
AGC TCC GTG CGG CTC AAG TGC GTG GCC AGC GGG CAC CCT CGG CCC GAC ATC ACG TGG ATG   611

 K  D  D  Q  A  L  T  R  P  E  A  A  E  P  R  K  K  K  W  T    205
AAG GAC GAC CAG GCC TTG ACG CGC CCA GAG GCC GCT GAG CCC AGG AAG AAG AAG TGG ACA   671

 L  S  L  K  N  L  R  P  E  D  S  G  K  Y  T  C  R  V  S  N    225
CTG AGC CTG AAG AAC CTG CGG CCG GAG GAC AGC GGC AAA TAC ACC TGC CGC GTG TCG AAC   731

 R  A  G  A  I  N  A  T  Y  K  V  D  V  I  Q  R  T  R  S  K    245
CGC GCG GGC GCC ATC AAC GCC ACC TAC AAG GTG GAT GTG ATC CAG CGG ACC CGT TCC AAG   791
```

FIG.4A

```
 P  V  L  T  G  T  H  P  V  N  T  T  V  D  F  G  G  T  T  S   265
CCC GTG CTC ACA GGC ACG CAC CCC GTG AAC ACG ACG GTG GAC TTC GGG GGG ACC ACG TCC   851

 F  Q  C  K  V  R  S  D  V  K  P  V  I  Q  W  L  K  R  V  E   285
TTC CAG TGC AAG GTG CGC AGC GAC GTG AAG CCG GTG ATC CAG TGG CTG AAG CGC GTG GAG   911

 Y  G  A  E  G  R  H  N  S  T  I  D  V  G  G  Q  K  F  V  V   305
TAC GGC GCC GAG GGC CGC CAC AAC TCC ACC ATC GAT GTG GGC GGC CAG AAG TTT GTG GTG   971

 L  P  T  G  D  V  W  S  R  P  D  G  S  Y  L  N  K  L  L  I   325
CTG CCC ACG GGT GAC GTG TGG TCG CGG CCC GAC GGC TCC TAC CTC AAT AAG CTG CTC ATC  1031

 T  R  A  R  Q  D  D  A  G  M  Y  I  C  L  G  A  N  T  M  G   345
ACC CGT GCC CGC CAG GAC GAT GCG GGC ATG TAC ATC TGC CTT GGC GCC AAC ACC ATG GGC  1091

 Y  S  F  R  S  A  F  L  T  V  L  P  D  P  K  P  P  G  P  P   365
TAC AGC TTC CGC AGC GCC TTC CTC ACC GTG CTG CCA GAC CCA AAA CCG CCA GGG CCA CCT  1151

 V  A  S  S  S  S  A  T  S  L  P  W  P  V  V  I  G  I  P  A   385
GTG GCC TCC TCG TCC TCG GCC ACT AGC CTG CCG TGG CCC GTG GTC ATC GGC ATC CCA GCC  1211

 G  A  V  F  I  L  G  T  L  L  L  W  L  C  Q  A  Q  K  K  P   405
GGC GCT GTC TTC ATC CTG GGC ACC CTG CTC CTG TGG CTT TGC CAG GCC CAG AAG AAG CCG  1271

 C  T  P  A  P  A  P  P  L  P  G  H  R  P  P  G  T  A  R  D   425
TGC ACC CCC GCG CCT GCC CCT CCC CTG CCT GGG CAC CGC CCG CCG GGG ACG GCC CGC GAC  1331

 R  S  G  D  K  D  L  P  S  L  A  A  L  S  A  G  P  G  V  G   445
CGC AGC GGA GAC AAG GAC CTT CCC TCG TTG GCC GCC CTC AGC GCT GGC CCT GGT GTG GGG  1391

 L  C  E  E  H  G  S  P  A  A  P  Q  H  L  L  G  P  G  P  V   465
CTG TGT GAG GAG CAT GGG TCT CCG GCA GCC CCC CAG CAC TTA CTG GGC CCA GGC CCA GTT  1451

 A  G  P  K  L  Y  P  K  L  Y  T  D  I  H  T  H  T  H  T  H   485
GCT GGC CCT AAG TTG TAC CCC AAA CTC TAC ACA GAC ATC CAC ACA CAC ACA CAC ACA CAC  1511

 S  H  T  H  S  H  V  E  G  K  V  H  Q  H  I  H  Y  Q  C  *   505
TCT CAC ACA CAC TCA CAC GTG GAG GGC AAG GTC CAC CAG CAC ATC CAC TAT CAG TGC TAG  1571

ACGGCACCGTATCTGCAGTGGGCACGGGGGGGCCGGCCAGACAGGCAGACTGGGAGGATGGAGGACGGAGCTGCAGACG  1650
```

FIG.4B

AAGGCAGGGGACCCATGGCGAGGAGGAATGGCCAGCACCCCAGGCAGTCTGTGTGTGAGGCATAGCCCCTGGACACACA 1729

CACACAGACACACACACTGCCTGGATGCATGTATGCACACACATGCGCGCACACGTGCTCCCTGAAGGCACACGTACGC 1808

ACACACGCACATGCACAGATATGCCGCCTGGGCACACAGATAAGCTGCCCAAATGCACGCACACGCACAGAGACATGCC 1887

AGAACATACAAGGACATGCTGCCTGAACATACACACGCACACCCATGCGCAGATGTGCTGCCTGGACACACACACACAC 1966

ACGGATATGCTGTCTGGACGCACACACGTGCAGATATGGTATCCGGACACACACGTGCACAGATATGCTGCCTGGACAC 2045

ACAGATAATGCTGCCTTGACACACACATGCACGGATATTGCCTGGACACACACACACACACGTGTGCACAGATATGCTG 2124

TCTGGACACGCACACACATGCAGATATGCTGCCTGGACACACACTTCCAGACACACGTGCACAGGCGCAGATATGCTGC 2203

CTGGACACACGCAGATATGCTGTCTAGTCACACACACACGCAGACATGCTGTCCGGACACACACACGCATGCACAGATA 2282

TGCTGTCCGGACACACACACGCACGCAGATATGCTGCCTGGACACACACACAGATAATGCTGCCTCAACACTCACACAC 2361

GTGCAGATATTGCCTGGACACACACATGTGCACAGATATGCTGTCTGGACATGCACACACGTGCAGATATGCTGTCCGG 2440

ATACACACGCACGCACACATGCAGATATGCTGCCTGGGCACACACTTCCGGACACACATGCACACACAGGTGCAGATAT 2519

GCTGCCTGGACACACGCAGACTGACGTGCTTTTGGGAGGGTGTGCCGTGAAGCCTGCAGTACGTGTGCCGTGAGGCTCA 2598

TAGTTGATGAGGGACTTTCCCTGCTCCACCGTCACTCCCCCAACTCTGCCCGCCTCTGTCCCCGCCTCAGTCCCCGCCT 2677

CCATCCCCGCCTCTGTCCCCTGGCCTTGGCGGCTATTTTTGCCACCTGCCTTGGGTGCCCAGGAGTCCCCTACTGCTGT 2756

GGGCTGGGGTTGGGGGCACAGCAGCCCCAAGCCTGAGAGGCTGGAGCCCATGGCTAGTGGCTCATCCCCACTGCATTCT 2835

CCCCCTGACACAGAGAAGGGGCCTTGGTATTTATATTTAAGAAATGAAGATAATATTAATAATGATGGAAGGAAGACTG 2914

GGTTGCAGGGACTGTGGTCTCTCCTGGGGCCCCGGACCCGCCTGGTCTTTCAGCCATGCTGATGACCACACCCCGTCCA 2993

GGCCAGACACCACCCCCCACCCCACTGTCGTGGTGGCCCCAGATCTCTGTAATTTTATGTAGAGTTTGAGCTGAAGCCC 3072

CGTATATTTAATTTATTTTGTTAAACATGAAAGTGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3151

AAAAAAAAGGGCGGCCGC 3169

FIG.4C

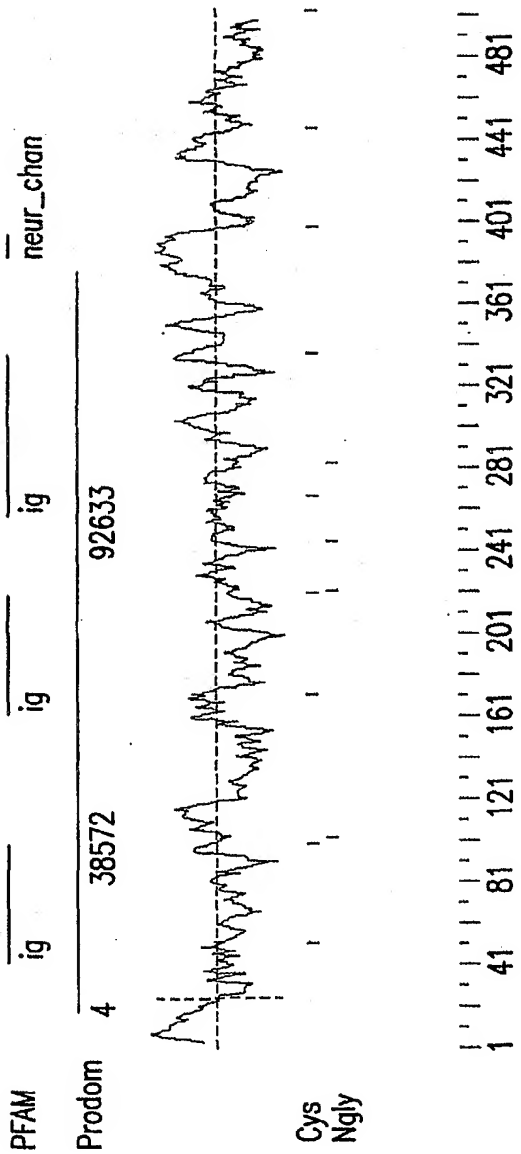


FIG.5

```
ig: domain 1 of 3, from 44 to 101: score 36.4, E = 9.9e-10

                *->GesvtLtCsvsgfgpp.p.vtWlrngk.............lslti.s
                   G +v+L+C v    g+p+p    W+++g++ +++ ++ +  +  l ++
      M003     44    GRTVRLQCPVE--GDPpPITMWTKDGRtihsgwsrfrvlpQGLKVkQ 88

                vtpeDsgGtYtCvv<-*
                v++eD+ G+Y C +
      M003     89 VEREDA-GVYVCKA     101

ig: domain 2 of 3, from 165 to 223: score 48.9, E = 1.3e-13

                *->GesvtLtCsvsgfgpp.p.vtWlrngk.............lslti.
                   G+sv+L C +s  g p+p++tW ++++  ++++  ++++++ +l ++
      M003    165    GSSVRLKCVAS--GHPrPdITWMKDDQaltrpeaaeprkkkWTLSLk 209

                svtpeDsgGtYtCvv<-*
                +++peDs G YtC+v
      M003    210 NLRPEDS-GKYTCRV     223

ig: domain 3 of 3, from 261 to 340: score 26.9, E = 8.8e-07

                *->GesvtLtCsvsgfgpp.p.vtWlrngk.................
                   G++ +++C v+  ++ +p ++Wl+  + + +++++++ + ++++
      M003    261    GGTTSFQCKVR--SDVkPvIQWLKRVEygaegrhnstidvggqkfvv 305

                ..............lslti.svtpeDsgGtYtCvv<-*
                 ++++  ++++++    l+i++++++D+ G Y C
      M003    306 lptgdvwsrpdgsyINKLLItRARQDDA-GMYICLG  . 340
```

FIG.6

neur_chan: domain 1 of 1, from 388 to 397: score 1.4, E = 9.7

```
                *->vfvIGTIgif<-*
                  vf+IGTI ++
MOO3    388     VFILGTLLLW     397
```

FIG.7

Input file M003jfmjf004c11; Output File M003jfmjf004c11.pat
Sequence length 1074

```
   R  V  R  P  T  G  D  V  W  S  R  P  D  G  S  Y  L  N  K    19
CA CGC GTC CGG CCC ACG GGT GAT GTG TGG TCA CGG CCT GAT GGC TCC TAC CTC AAC AAG    59

 L  L  I  S  R  A  R  Q  D  D  A  G  M  Y  I  C  L  G  A  N    39
CTG CTC ATC TCT CGG GCC CGC CAG GAT GAT GCT GGC ATG TAC ATC TGC CTA GGT GCA AAT   119

 T  M  G  Y  S  F  R  S  A  F  L  T  V  L  P  D  P  K  P  P    59
ACC ATG GGC TAC AGT TTC CGT AGC GCC TTC CTC ACT GTA TTA CCA GAC CCC AAA CCT CCA   179

 G  P  P  M  A  S  S  S  S  S  T  S  L  P  W  P  V  V  I  G    79
GGG CCT CCT ATG GCT TCT TCA TCG TCA TCC ACA AGC CTG CCA TGG CCT GTG GTG ATC GGC   239

 I  P  A  G  A  V  F  I  L  G  T  V  L  L  W  L  C  Q  T  K    99
ATC CCA GCT GGT GCT GTC TTC ATC CTA GGC ACT GTG CTG CTC TGG CTT TGC CAG ACC AAG   299

 K  K  P  C  A  P  A  S  T  L  P  V  P  G  H  R  P  P  G  T   119
AAG AAG CCA TGT GCC CCA GCA TCT ACA CTT CCT GTG CCT GGG CAT CGT CCC CCA GGG ACA   359

 S  R  E  R  S  G  D  K  D  L  P  S  L  A  V  G  I  C  E  E   139
TCC CGA GAA CGC AGT GGT GAC AAG GAC CTG CCC TCA TTG GCT GTG GGC ATA TGT GAG GAG   419

 H  G  S  A  M  A  P  Q  H  I  L  A  S  G  S  T  A  G  P  K   159
CAT GGA TCC GCC ATG GCC CCC CAG CAC ATC CTG GCC TCT GGC TCA ACT GCT GGC CCC AAG   479

 L  Y  P  K  L  Y  T  D  V  H  T  H  T  H  T  H  T  C  T  H   179
CTG TAC CCC AAG CTA TAC ACA GAT GTG CAC ACA CAC ACA CAT ACA CAC ACC TGC ACT CAC   539

 T  L  S  C  W  R  A  R  F  I  N  T  S  M  S  T  I  S  A  K   199
ACG CTC TCA TGT TGG AGG GCA AGG TTC ATC AAC ACC AGC ATG TCC ACT ATC AGT GCT AAA   599

 Y  S  E  S  P  S  T  V  S  *                                 209
TAC AGC GAA TCT CCA AGC ACT GTG TCC TGA                       629
```

FIG.8A

GGTAGGCATTTGGGGGCCAAGGCAACAGGTTGGGAGAATTGAGAACAATGGAGGAAGAGTATCTTAGGGTGCCTTATGG 708

TGGACACTCACAAACTTGGCCATATAGATGTATGTACTACCAGATGAACAGCCAGCCAGATTCACACACGCACATGTTT 787

AAACGTGTAAACGTGTGCACAACTGCACACACAACCTGAGAAACCTTCAGGAGGATTTGTGGTGTGACTTTGCAGTGAC 866

ATGTAGCGATGGCTAGTTGAAGGAATCTCCCTCATGTCTTAGTGGTCATGGCCACTTCCCCACCCCTGCCCATCTGTGT 945

TCCTGCCTGGCCTTGGTGGTGCTTCCGTGTGCCCTGGGTTTTCCAGGAACCCTATCAACCTGACTGGGGTGAGCAGTGC 1024

AGCCATGCNTGGAGGTTTGAGCCACCCTCCCCTTGCTAGAGAGAAGGGCN 1074

FIG.8B

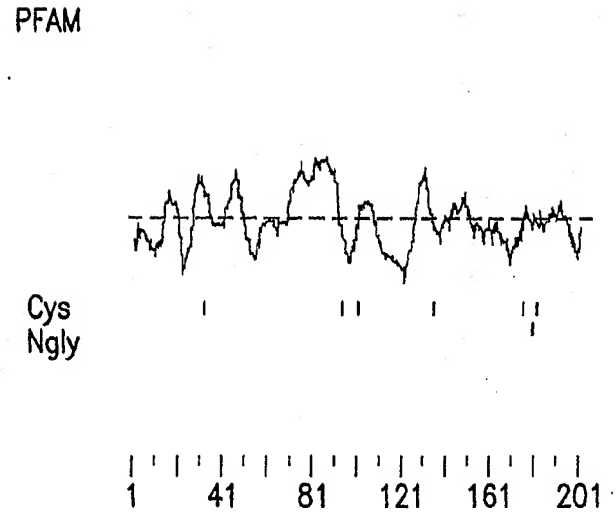


FIG.9

Input file M347Alhbad295g12; Output File M347Alhbad295g12.pat
Sequence length 1423

```
                                        M  P  G  P  R  V  W  G  K  Y  L  W    12
GTCGACCCACGCGTCCGCCCACGCGTCCGG ATG CCT GGA CCC AGA GTG TGG GGG AAA TAT CTC TGG   66

 R  S  P  H  S  K  G  C  P  G  A  M  W  W  L  L  L  W  G  V    32
AGA AGC CCT CAC TCC AAA GGC TGT CCA GGC GCA ATG TGG TGG CTG CTT CTC TGG GGA GTC  126

 L  Q  A  C  P  T  R  G  S  V  L  L  A  Q  E  L  P  Q  Q  L    52
CTC CAG GCT TGC CCA ACC CGG GGC TCC GTC CTC TTG GCC CAA GAG CTA CCC CAG CAG CTG  186

 T  S  P  G  Y  P  E  P  Y  G  K  G  Q  E  S  S  T  D  I  K    72
ACA TCC CCC GGG TAC CCA GAG CCG TAT GGC AAA GGC CAA GAG AGC AGC ACG GAC ATC AAG  246

 A  P  E  G  F  A  V  R  L  V  F  Q  D  F  D  L  E  P  S  Q    92
GCT CCA GAG GGC TTT GCT GTG AGG CTC GTC TTC CAG GAC TTC GAC CTG GAG CCG TCC CAG  306

 D  C  A  G  D  S  V  T  V  S  W  G  W  G  G  S  R  Q  D  C   112
GAC TGT GCA GGG GAC TCT GTC ACA GTG AGC TGG GGA TGG GGG GGG TCC CGC CAG GAC TGT  366

 G  Q  G  D  S  R  G  C  G  K  W  R  C  P  E  S  P  I  W  R   132
GGC CAG GGA GAT TCC CGG GGT TGT GGG AAG TGG CGG TGC CCT GAA TCC CCC ATC TGG AGG  426

 R  D  E  F  S  M  *                                           139
AGG GAT GAA TTT TCC ATG TAG                                    447

GGGCAGTCGGGCTTGGCTTACCGGGGAGCAGTGGTGGACCCCAGGACACAGCCTCCCACCAGCGCCTCCGGGGCTGCCA  526

TCTGGGCCCCACAGAGCAAAGAGGGCAGCAAGCAGGCCCTGCGTTTGGAAGGCTTATGAATGGACACACAAATCTTGCA  605

AATCTATGGAGCCAGGGGCAGGGACGCACATATTGGTTGTTAAAAATATGTCATCATGTATTTGTTGAGTGCCTGCTCT  684

ATCAGGTGAGGAAGCTGGACACAAATAATAACAAAAGATTAAGTCACCGTTCACACTTACCTTGGAAGAGCTATTACAA  763

AACTTCTAACGCCAAAGCCTTATTCAGAATAAGGACATTTTAAAAACAGTACTTGATGGAGTGATGCAAGCTTGCAGTC  842

CCAGCAGTATAGTCAGGAGACTGAGGCTGGAGGATCAGAGGGCTGGAGCCCAGGGTTCAAGGCCAGCCTAAGCAACATA  921
```

FIG.10A

GCAAGACCCCATCTCAAAAATAAGTAAATAATAAATAAAAATAAAAAGAGCACATTATCTTTTGATTTAAATTTTATTT 1000

ATATCAAAATGACATAAATTTTTGAACTTTATTTTTTAATTTTAAAATTTTTAATTATTATGGATACATAATAGTTGTA 1079

AGACTTTTTGTTTTTTAATTAAAGTTTTCTAAGGCTGGGCGCAGTAGCTCATGTCTGTAGTCCCAGCACTTTGGGAGGC 1158

TGAGGCGAAAGAAGCACTTGAGCCCAGGAATTTGAGACCAGCCTGGGCAACATAGCAAGACCCCATCTCTACAAAAAAA 1237

TTTAAAAATTAGCCAAGTGTGGTGGCACGCACCTGTGGTCCCAGCTACAAGGGACGCTGAAGTGAGAGGATCACTTGAG 1316

CCTGGAAGGTAGAGGCTGCAGTGAGCTCTGATCATGACACCGTACTCCAGCCTGGGTGACAGAGTGAGACCCTGTCTCC 1395

AAAAAAAAAAAAAAAAAAGGGCGGCCGC 1423

FIG.10B

M347

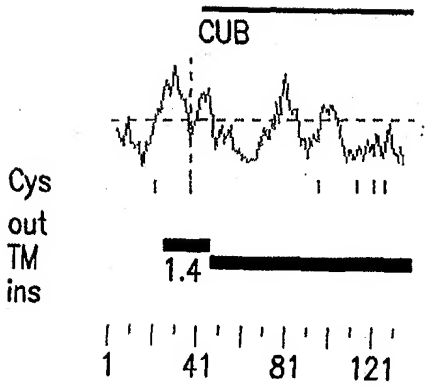


FIG.11

```
CUB: domain 1 of 1, from 40 to 136: score -17.7, E = 0.035

                *->CGgtIdItessGsisSPnYPnrsdYppnkeCvWrIrappgyrvVeLt
                   G +I+ +e +  ++SP+YP+  +Y   +e    I ap+g+  V L
hM347     40    -GSVLLAQELPQQLTSPGYPE--PYGKGQESSTDIKAPEGFA-VRLV 82

                FqdFdIEdhdgapCryDyvEirDGdpss.pIIG....rfCG....sgkPe
                FqdFdIE +++  C+ D+v +  G ++s++  G++++r CG+ +  ++P
hM347     83 FQDFDLEPSQD—CAGDSVTVSWGWGGSrQDCGqgdsRGCGkwrcPESP- 129

                dirStsnrmlikFvsDasvskrGFkAty<-*
                          +  +D+          +
hM347    130-----------IWRRDE----------F    136
```

FIG. 12

Input file T272Athda89h3; Output File T272Athda89h3.pat
Sequence length 5036

```
GTCGACCCACGCGTCCGCTCGAAGCGGGGACCCTCGCCCCGTCCTCGGCTGTCCAGTCCTCCTCCTCGCAGACCCCGGC  79

GGTTCCTACCCCAGGCCGCAGGGGAGACGGTGCCCCAAGGCAGGCTTCATATCCTGAACGCTGGGATCCCCCAGGACAT 158

                                                                      M  S     2
TCCCTGGCCCCCAGGCCCCAGGTCCCAGGCCCCAGGGCTGAGCTGTGGGCAGGCCCCACCTGGCCTCTGCA ATG TCA  235

 P  P  L  C  P  L  L  L  L  A  V  G  L  R  L  A  G  T  L  N    22
CCG CCT CTG TGT CCC CTC CTT CTC CTG GCT GTG GGC CTG CGG CTG GCT GGA ACT CTC AAC  295

 P  S  D  P  N  T  C  S  F  W  E  S  F  T  T  T  T  K  E  S    42
CCC AGT GAT CCC AAT ACC TGC AGC TTC TGG GAA AGC TTC ACT ACC ACC ACC AAG GAG TCC  355

 H  S  R  P  F  S  L  L  P  S  E  P  C  E  R  P  W  E  G  P    62
CAC TCC CGC CCC TTC AGC CTG CTC CCC TCA GAG CCC TGC GAG CGG CCC TGG GAG GGC CCC  415

 H  T  C  P  S  P  Q  T  Q  R  K  L  L  A  S  R  D  S  F  C    82
CAT ACT TGC CCC AGC CCA CAA ACT CAG AGG AAA CTC CTG GCT TCT AGG GAT TCA TTC TGC  475

 M  V  C  V  G  A  G  V  Q  W  R  D  R  S  A  L  Q  P  Q  T   102
ATG GTC TGT GTC GGG GCT GGA GTG CAG TGG CGA GAT CGT AGT GCA CTG CAA CCT CAA ACA  535

 G  N  A  L  S  M  R  P  Q  P  R  V  L  S  G  A  P  S  L  A   122
GGG AAT GCG CTT TCT ATG CGC CCT CAG CCC AGA GTG TTG AGT GGT GCC CCT TCC CTG GCC  595

 S  P  G  H  T  V  V  V  K  T  D  H  R  Q  R  L  Q  C  C  H   142
TCC CCT GGC CAC ACT GTG GTG GTG AAG ACG GAC CAC CGC CAG CGC CTG CAG TGC TGC CAT  655

 G  F  Y  E  S  R  G  F  C  V  P  L  C  A  Q  E  C  V  H  G   162
GGC TTC TAT GAG AGC AGG GGG TTC TGT GTC CCG CTC TGT GCC CAG GAG TGT GTC CAT GGC  715

 R  C  V  A  P  N  Q  C  Q  C  V  P  G  W  R  G  D  D  C  S   182
CGT TGT GTG GCA CCC AAT CAG TGC CAA TGT GTG CCA GGC TGG CGG GGC GAC GAC TGT TCC  775

 S  A  P  N  C  L  Q  P  C  T  P  G  Y  Y  G  P  A  C  Q  F   202
AGT GCC CCG AAC TGC CTT CAG CCC TGT ACC CCT GGC TAC TAT GGC CCT GCC TGC CAG TTC  835
```

FIG.13A

```
 R  C  Q  C  H  G  A  P  C  D  P  Q  T  G  A  C  F  C  P  A   222
CGC TGC CAG TGC CAT GGG GCA CCC TGC GAT CCC CAG ACT GGA GCC TGC TTC TGC CCC GCA  895

 E  R  T  G  P  S  C  D  V  S  C  S  Q  G  T  S  G  F  F  C   242
GAG AGA ACT GGG CCC AGC TGT GAC GTG TCC TGT TCC CAG GGC ACT TCT GGC TTC TTC TGC  955

 P  S  T  H  P  C  Q  N  G  G  V  F  Q  T  P  Q  G  S  C  S   262
CCC AGC ACC CAT CCT TGC CAA AAT GGA GGT GTC TTC CAA ACC CCA CAG GGC TCC TGC AGC 1015

 C  P  P  G  W  M  G  T  I  C  S  L  P  C  P  E  G  F  H  G   282
TGC CCC CCT GGC TGG ATG GGC ACC ATC TGC TCC CTG CCC TGC CCA GAG GGC TTT CAC GGA 1075

 P  N  C  S  Q  E  C  R  C  H  N  G  G  L  C  D  R  F  T  G   302
CCC AAC TGC TCC CAG GAA TGT CGC TGC CAC AAC GGC GGC CTC TGT GAC CGA TTC ACT GGG 1135

 Q  C  R  C  A  P  G  Y  T  G  D  R  C  R  E  E  C  P  V  G   322
CAG TGC CGC TGC GCT CCG GGT TAC ACT GGG GAT CGG TGC CGG GAG GAG TGC CCG GTG GGC 1195

 R  F  G  Q  D  C  A  E  T  C  D  C  A  P  D  A  R  C  F  P   342
CGC TTT GGG CAG GAC TGT GCT GAG ACG TGC GAC TGC GCC CCG GAC GCC CGT TGC TTC CCG 1255

 A  N  G  A  C  L  C  E  H  G  F  T  G  D  R  C  T  D  R  L   362
GCC AAC GGC GCA TGT CTG TGC GAA CAC GGC TTC ACT GGG GAC CGC TGC ACG GAT CGC CTC 1315

 C  P  D  G  F  Y  G  L  S  C  Q  A  P  C  T  C  D  R  E  H   382
TGC CCC GAC GGC TTC TAC GGT CTC AGC TGC CAG GCC CCC TGC ACC TGC GAC CGG GAG CAC 1375

 S  L  S  C  H  P  M  N  G  E  C  S  C  L  P  G  W  A  G  L   402
AGC CTC AGC TGC CAC CCG ATG AAC GGG GAG TGC TCC TGC CTG CCG GGC TGG GCG GGC CTC 1435

 H  C  N  E  S  C  P  Q  D  T  H  G  P  G  C  Q  E  H  C  L   422
CAC TGC AAC GAG AGC TGC CCG CAG GAC ACG CAT GGG CCA GGG TGC CAG GAG CAC TGT CTC 1495

 C  L  H  G  G  V  C  Q  A  T  S  G  L  C  Q  C  A  P  G  Y   442
TGC CTG CAC GGT GGC GTC TGC CAG GCT ACC AGC GGC CTC TGT CAG TGC GCG CCG GGT TAC 1555

 T  G  P  H  C  A  S  L  C  P  P  D  T  Y  G  V  N  C  S  A   462
ACG GGC CCT CAC TGT GCT AGT CTT TGT CCT CCT GAC ACC TAC GGT GTC AAC TGT TCT GCA 1615
```

FIG.13B

```
 R  C  S  C  E  N  A  I  A  C  S  P  I  D  G  E  C  V  C  K   482
CGC TGC TCA TGT GAA AAT GCC ATC GCC TGC TCA CCC ATC GAC GGC GAG TGC GTC TGC AAG  1675

 E  G  W  Q  R  G  N  C  S  V  P  C  P  P  G  T  W  G  F  S   502
GAA GGT TGG CAG CGT GGT AAC TGC TCT GTG CCC TGC CCA CCC GGA ACC TGG GGC TTC AGT  1735

 C  N  A  S  C  Q  C  A  H  E  A  V  C  S  P  Q  T  G  A  C   522
TGC AAT GCC AGC TGC CAG TGT GCC CAT GAG GCA GTC TGC AGC CCC CAA ACT GGA GCC TGT  1795

 T  C  T  P  G  W  H  G  A  H  C  Q  L  P  C  P  K  G  Q  F   542
ACC TGC ACC CCT GGG TGG CAT GGG GCC CAC TGC CAG CTG CCC TGT CCG AAG GGG CAG TTT  1855

 G  E  G  C  A  S  R  C  D  C  D  H  S  D  G  C  D  P  V  H   562
GGA GAA GGT TGT GCC AGT CGC TGT GAC TGT GAC CAC TCT GAT GGC TGT GAC CCT GTT CAT  1915

 G  R  C  Q  C  Q  A  G  W  M  G  A  R  C  H  L  S  C  P  E   582
GGA CGC TGT CAG TGC CAG GCT GGC TGG ATG GGT GCC CGC TGC CAC CTG TCC TGC CCT GAG  1975

 G  L  W  G  V  N  C  S  N  T  C  T  C  K  N  G  G  T  C  L   602
GGC TTA TGG GGA GTC AAC TGT AGC AAC ACC TGC ACC TGC AAG AAT GGG GGC ACC TGT CTC  2035

 P  E  N  G  N  C  V  C  A  P  G  F  R  G  P  S  C  Q  R  S   622
CCT GAG AAT GGC AAC TGC GTG TGT GCA CCC GGA TTC CGG GGC CCC TCC TGC CAG AGA TCC  2095

 C  Q  P  G  R  Y  G  K  R  C  V  P  C  K  C  A  N  H  S  F   642
TGT CAG CCT GGC CGC TAT GGC AAA CGC TGT GTG CCC TGC AAG TGC GCT AAC CAC TCC TTC  2155

 C  H  P  S  N  G  T  C  Y  C  L  A  G  W  T  G  P  D  C  S   662
TGC CAC CCC TCG AAC GGG ACC TGC TAC TGC CTG GCT GGC TGG ACA GGC CCC GAC TGC TCC  2215

 Q  P  C  P  P  G  H  W  G  E  N  C  A  Q  T  C  Q  C  H  H   682
CAG CCA TGC CCT CCA GGA CAC TGG GGA GAA AAC TGT GCC CAG ACC TGC CAA TGT CAC CAT  2275

 G  G  T  C  H  P  Q  D  G  S  C  I  C  P  L  G  W  T  G  H   702
GGT GGG ACC TGC CAT CCC CAG GAT GGG AGC TGT ATC TGC CCC CTA GGC TGG ACT GGA CAC  2335

 H  C  L  E  G  C  P  L  G  T  F  G  A  N  C  S  Q  P  C  Q   722
CAC TGC TTA GAA GGC TGC CCT CTG GGG ACA TTT GGT GCT AAC TGC TCC CAG CCA TGC CAG  2395
```

FIG.13C

```
 C  G  P  G  E  K  C  H  P  E  T  G  A  C  V  C  P  P  G  H   742
TGT GGT CCT GGA GAA AAG TGC CAC CCA GAG ACT GGG GCC TGT GTA TGT CCC CCA GGG CAC  2455

 S  G  A  P  C  R  I  G  I  Q  E  P  F  T  V  M  P  T  T  P   762
AGT GGT GCA CCT TGC AGG ATT GGA ATC CAG GAG CCC TTT ACT GTG ATG CCG ACC ACT CCA  2515

 V  A  Y  N  S  L  G  A  V  I  G  I  A  V  L  G  S  L  V  V   782
GTA GCG TAT AAC TCG CTG GGT GCA GTG ATT GGC ATT GCA GTG CTG GGG TCC CTT GTG GTA  2575

 A  L  V  A  L  F  I  G  Y  R  H  W  Q  K  G  K  E  H  H  H   802
GCC CTG GTG GCA CTG TTC ATT GGC TAT CGG CAC TGG CAA AAA GGC AAG GAG CAC CAC CAC  2635

 L  A  V  A  Y  S  S  G  R  L  D  G  S  E  Y  V  M  P  D  V   822
CTG GCT GTG GCT TAC AGC AGC GGG CGC CTG GAC GGC TCC GAG TAT GTC ATG CCA GAT GTC  2695

 P  P  S  Y  S  H  Y  Y  S  N  P  S  Y  H  T  L  S  Q  C  S   842
CCT CCG AGC TAC AGT CAC TAC TAC TCC AAC CCC AGC TAC CAC ACC CTG TCG CAG TGC TCC  2755

 P  N  P  P  P  P  N  K  V  P  G  P  L  F  A  S  L  Q  N  P   862
CCA AAC CCC CCA CCC CCT AAC AAG GTT CCA GGC CCG CTC TTT GCC AGC CTG CAG AAC CCT  2815

 E  R  P  G  G  A  Q  G  H  D  N  H  T  T  L  P  A  D  W  K   882
GAG CGG CCA GGT GGG GCC CAA GGG CAT GAT AAC CAC ACC ACC CTG CCT GCT GAC TGG AAG  2875

 H  R  R  E  P  P  P  G  P  L  D  R  G  S  S  R  L  D  R  S   902
CAC CGC CGG GAG CCC CCT CCA GGG CCT CTG GAC AGG GGG AGC AGC CGC CTG GAC CGA AGC  2935

 Y  S  Y  S  Y  S  N  G  P  G  P  F  Y  D  K  G  L  I  S  E   922
TAC AGC TAT AGC TAC AGC AAT GGC CCA GGC CCA TTC TAC GAT AAA GGG CTC ATC TCT GAA  2995

 E  E  L  G  A  S  V  A  S  L  S  S  E  N  P  Y  A  T  I  R   942
GAG GAG CTC GGG GCC AGT GTG GCT TCC CTG AGC AGT GAG AAC CCA TAT GCC ACC ATC CGG  3055

 D  L  P  S  L  P  G  G  P  R  E  S  S  Y  M  E  M  K  G  P   962
GAC CTG CCC AGC TTG CCA GGG GGC CCC CGG GAG AGC AGC TAC ATG GAG ATG AAA GGC CCT  3115

 P  S  G  S  A  P  R  Q  P  P  Q  F  W  D  S  Q  R  R  R  Q   982
CCC TCA GGA TCT GCC CCC AGG CAG CCT CCT CAG TTT TGG GAC AGC CAG AGG CGG CGG CAA  3175
```

FIG.13D

```
 P  Q  P  Q  R  D  S  G  T  Y  E  Q  P  S  P  L  I  H  D  R  1002
CCC CAG CCA CAG AGA GAC AGT GGC ACC TAC GAG CAG CCC AGC CCC CTG ATC CAT GAC CGA  3235

 D  S  V  G  S  Q  P  P  L  P  P  G  L  P  P  G  H  Y  D  S  1022
GAC TCT GTG GGC TCC CAG CCC CCT CTG CCT CCG GGC CTA CCC CCC GGC CAC TAT GAC TCA  3295

 P  K  N  S  H  I  P  G  H  Y  D  L  P  P  V  R  H  P  P  S·  1042
CCC AAG AAC AGC CAC ATC CCT GGA CAT TAT GAC TTG CCT CCA GTA CGG CAT CCC CCA TCA  3355

 P  P  L  R  R  Q  D  R  *                                    1051
CCT CCA CTT CGA CGC CAG GAC CGT TGA                           3382

GGAGCCAGGATGGTATGGCAGAGGCCAGCACACCTGGCTGTTGCTGCTCAAGGCTGGGGACAGAGCCTAGTGTACCCCT  3461

GCCAGGAGCAGGGAGTGGACCGGCAGGCTGTGAACATGAACAACGCTTAACAGAGCAAGTGATGGGAGCCTTGTTCCTG  3540

GGTTCTACCATGGGAGACGCTGATCAGCAGGATGCCTGGCTCCCTTTCCCAACCCACTGCTCCCAAGGCCTCCAGGGCC  3619

CTGTGTACATAAACTGGTGGGTTGGAAGTTGCTGGGTAACTCTGATTTCAGACATGCGTGTGGGGTACCTTTTCTGTGC  3698

ATGCTCAGCCTGGGCTCTGTGCGTGTGTGTTTCTGTGATTTTAGAAGGGTACCAGGCAGGTTCTGTCCTAGGGCACT  3777

TACCATTTAGTAGGGAGATGGAACCAACCCAATTAACTCTAGCAATAGCCTCCTAACTGGCCTCCTCCATTGATTCAGT  3856

GAACCTTCCAATGCATGGCTCATAATTTCAAAATACAGGCTGGTTAGTTACTCCCTACCTGAAAGCCTTCATAGGTGCC  3935

TCTTTGCTCTTCTGCCAGTATCAAAACTTTTGAAGGCCTTAAAGGCCCTGCTTTGCCTGGCCCATCTGTCTCTCCAGCC  4014

TCACCTTGAACTGTGTTCCTGTCACTGCACGCCAGTCACACCGGCCTCTAGGTCCTCCTGTAGGCCACTCTTCTTTCTG  4093

GCACAGGGACCTGCACACCTGGAGTGCCCTTCCTCCCCCACTCGCCTGTTCACCCCTGCTTTTCCTTTACACCTCCTCC  4172

TCAGGGAAGTGCCCACCCTCCGTACATCTTTCACAGCCCTGATTGCAGCTGTGTTCACTCACCAGGTACCTGCAGAAGG  4251

CCTACAGGGTGCCAGGCACTTCTTTAATGGGTTCTTTCTTTATGTGATTATTTGATTAATCTCTGCCTCCCCCACTAGA  4330

CTGTAAGCTCCCTGAAGGCAAGAATCCTGTGCTTATGCTCAATATTAGCTCTCCCTTGGCACAGAGTAGGCACTCAACA  4409

AATGCTCCCCAAAAGGCTGAGTGGCTGACTGAATTAAGTACCAGTGACATGCAGTAACTGCTAAGATAGATGAGCCATC  4488
```

FIG.13E

TGTATGCTCTGACAGTTACAGACTGAATAAGTTGGAGACTTCCCTAAAGGGTGGCATTTCCCCAGGGTAACAACGCAGA 4567

GCTCAGGTGTGGGAAGGTGCCAGGGGCAGGGGTGCAGAGGGGCTGAGGCTGAGGGGGGTGCAGAGGCTGGAGAAAGGAT 4646

AACAGGAGAGAGTATACAGGCATGCCTTGATTTATTGCACTTCACAGGTAGCAGAATTTTTAAAGAAATTGAAGGTTTT 4725

GGGACATATATGTGACAGCAATAGGTTAAGAAAAGCAAAGCAGAGAAATTGAAGATTTGTGTCAACACTGCTTTAAGCA 4804

AATCTGTTGGCACCATTTTTCCAATAGCATGTGCCCATTTTGGGTCTCTACATTGCATTTTGGTAATTGCTTGCAATAT 4883

TTCAAGCATTTTCATTGTTATTATATGTGTTATAGTGATCTGTGATCAGTGATCTTTGATATATTATTGTAATTGTTTC 4962

GGGGCGCCATGAACCGCACCCATATAACACGGTAAACTTAATCAGCAAAAAAAAAAAAAAAAAAGGGCGGCCG 5036

FIG.13F

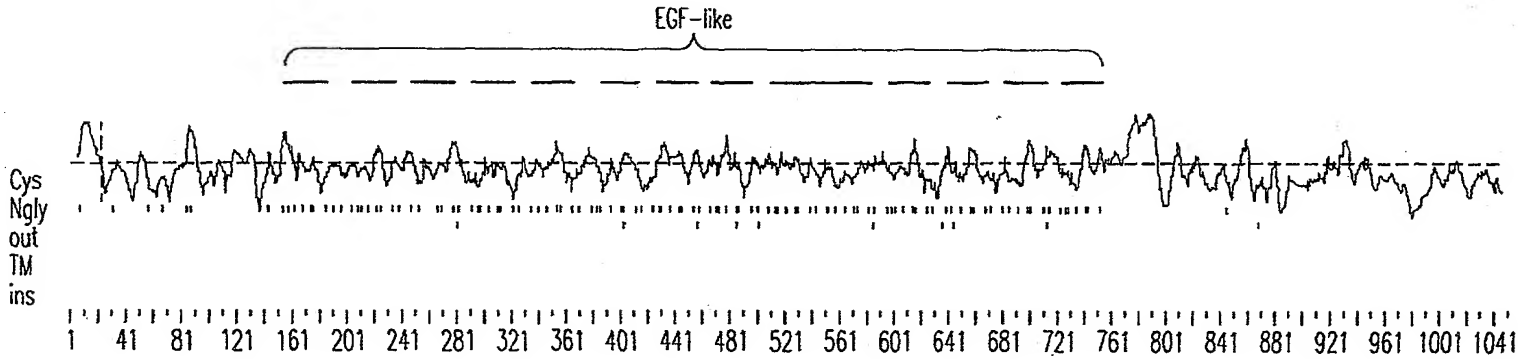


FIG. 14

EGF: domain 1 of 14, from 151 to 181: score 14.0, E = 1.2

```
                *->Capnn..pCsngGtCvntpggssdnfggytCeCppGdyylsytGkrC
                  C p++ + C + G+Cv              +C+C pG     + G++C
hT272    151      CVPLCaqECVH-GRCVAPN----------QCQCVPG-----WRGDDC 181
```

EGF: domain 2 of 14, from 200 to 229: score -2.2, E = 36

```
                *->CapnnpCsngGtCvntpggssdnfggytCeCppGdyylsytGkrC<-
                  C+  + C++ + C +  g          C+Cp         tG+ C
hT272    200      CQFRCQCHG-APCDPQTG---------ACFCPAE-----RTGPSC   229
```

EGF: domain 3 of 14, from 242 to 272: score 16.0, E = 0.81

```
                *->CapnnpCsngGtCvntpggssdnfggytCeCppGdyylsytGkrC<-
                  C+++ pC+ngG+ +   g          +C CppG     + G  C
hT272    242      CPSTHPCQNGGVFQTPQG---------SCSCPPG-----WMGTIC   272
```

EGF: domain 4 of 14, from 285 to 315: score 27.0, E = 0.00045

```
                *->CapnnpCsngGtCvntpggssdnfggytCeCppGdyylsytGkrC<-
                  C++++ C+ngG C    g          +C+C+pG     ytG+rC
hT272    285      CSQECRCHNGGLCDRFTG---------QCRCAPG-----YTGDRC   315
```

FIG. 15A

EGF: domain 5 of 14, from 328 to 358: score 18.0, E = 0.22

```
                *->CapnnpCsngGtCvntpggssdnfggytCeCppGdyylsytGkrC<-
                  Ca+++ C  +++C +  g          C C +G     +tG+rC
hT272    328    CAETCDCAPDARCFPANG---------ACLCEHG-----FTGDRC    358
```

EGF: domain 6 of 14, from 378 to 404: score 7.4, E = 4.9

```
                *->CapnnpCsngGtCvntpggssdnfggytCeCppGdyylsytGkrC<-
                  C+ +    +   C++  g         +C C pG     ++G +C
hT272    378    CDRE----HSLSCHPMNG---------ECSCLPG-----WAGLHC    404
```

EGF: domain 7 of 14, from 417 to 447: score 29.2, E = 9.3e-05

```
                *->CapnnpCsngGtCvntpggssdnfggytCeCppGdyylsytGkrC<-
                  C++++ C++gG+C+ t g          C+C+pG     ytG++C
hT272    417    CQEHCLCLHGGVCQATSG---------LCQCAPG-----YTGPHC    447
```

EGF: domain 8 of 14, from 460 to 490: score 6.0, E = 6.5

```
                *->CapnnpCsngGtCvntpggssdnfggytCeCppGdyylsytGkrC<-
                  C+  + C n   C +  g         +C+C++G     ++  +C
hT272    460    CSARCSCENAIACSPIDG---------ECVCKEG-----WQRGNC    490
```

FIG.15B

EGF: domain 9 of 14, from 503 to 533: score 15.9, E = 0.82

```
                *->CapnnpCsngGtCvntpggssdnfggytCeCppGdyylsytGkrC<-
                  C+  + C + ++C +  g          C+C+pG     ++G +C
hT272   503     CNASCQCAHEAVCSPQTG---------ACTCTPG-----WHGAHC   533
```

DSL: domain 1 of 1, from 518 to 576: score -20.5, E = 6.8

```
                *->WstdkhiggrtslGfnleyrirvtCdenYYGegCnkFCrPrdDafgH
                   +t + + + +      + +    C +  +GegC+  C+       H
hT272   518     -QTGACTCTPG----WHGAHCQLPCPKGQFGEGCASRCDCD-----H 554

                yt.Cd.enGnklCleGWkGeyC<-*
                + +Cd+ +G+ +C +GW+G  C
hT272   555 SDgCDpVHGRCQCQAGWMGARC    576
```

EGF: domain 10 of 14, from 546 to 576: score 11.7, E = 2

```
                *->CapnnpCsngGtCvntpggssdnfggytCeCppGdyylsytGkrC<-
                  Ca+ + C++   C +++g         +C+C+ G     + G rC
hT272   546     CASRCDCDHSDGCDPVHG---------RCQCQAG-----WMGARC   576
```

EGF: domain 11 of 14, from 589 to 619: score 17.9, E = 0.24

```
                *->CapnnpCsngGtCvntpggssdnfggytCeCppGdyylsytGkrC<-
                  C+ ++ C+ngGtC++  g          C+C+pG     + G+ C
hT272   589     CSNTCTCKNGGTCLPENG---------NCVCAPG-----FRGPSC   619
```

FIG.15C

EGF: domain 12 of 14, from 632 to 661: score 18.0, E = 0.23

```
                  *->CapnnpCsngGtCvntpggssdnfggytCeCppGdyylsytGkrC<-
                    C p   C n+ +C+++ g          tC C  G     +tG++C
hT272    632      CVPC-KCANHSFCHPSNG---------TCYCLAG-----WTGPDC    661
```

EGF: domain 13 of 14, from 674 to 704: score 27.1, E = 0.00042

```
                  *->CapnnpCsngGtCvntpggssdnfggytCeCppGdyylsytGkrC<-
                    Ca+++ C++gGtC++  g         +C+Cp G     +tG++C
hT272    674      CAQTCQCHHGGTCHPQDG---------SCICPLG-----WTGHHC    704
```

EGF: domain 14 of 14, from 717 to 747: score 1.7, E = 16

```
                  *->CapnnpCsngGtCvntpggssdnfggytCeCppGdyylsytGkrC<-
                    C++++ C  g +C++  g          C+CppG      +G  C
hT272    717      CSQPCQCGPGEKCHPETG---------ACVCPPG-----HSGAPC    747
```

FIG.15D

Input file t272Atmzb62c4; Output File t272Atmzb62c4.pat
Sequence length 2569

```
   S   T   H   A   S   G   D   P   V   H   G   Q   C   R   C   Q   A   G   W      19
G TCG ACC CAC GCG TCC GGT GAC CCT GTT CAT GGA CAG TGC CGA TGT CAG GCT GGT TGG      58

 M   G   T   R   C   H   L   P   C   P   E   G   F   W   G   A   N   C   S   N    39
ATG GGC ACA CGC TGC CAC CTG CCT TGC CCG GAG GGC TTT TGG GGA GCC AAC TGC AGT AAC   118

 T   C   T   C   K   N   G   G   T   C   V   S   E   N   G   N   C   V   C   A    59
ACC TGT ACC TGC AAG AAT GGT GGT ACC TGT GTG TCT GAG AAT GGC AAC TGC GTG TGC GCA   178

 P   G   F   R   G   P   S   C   Q   R   P   C   P   P   G   R   Y   G   K   R    79
CCA GGG TTC CGA GGC CCC TCC TGC CAG AGG CCC TGC CCG CCT GGT CGC TAT GGC AAA CGC   238

 C   V   Q   C   K   C   N   N   N   H   S   S   C   H   P   S   D   G   T   C    99
TGT GTG CAA TGC AAG TGT AAC AAC AAC CAT TCT TCC TGC CAC CCA TCG GAC GGG ACC TGC   298

 S   C   L   A   G   W   T   G   P   D   C   S   E   A   C   P   P   G   H   W   119
TCC TGC CTG GCG GGC TGG ACA GGC CCT GAC TGC TCC GAG GCA TGT CCC CCA GGC CAC TGG   358

 G   L   K   C   S   Q   L   C   Q   C   H   H   G   G   T   C   H   P   Q   D   139
GGA CTC AAA TGC TCC CAA CTC TGC CAG TGT CAT CAT GGT GGG ACC TGC CAC CCC CAG GAT   418

 G   S   C   I   C   T   P   G   W   T   G   P   N   C   L   E   G   C   P   P   159
GGG AGC TGT ATC TGC ACG CCA GGC TGG ACT GGA CCC AAC TGC TTG GAA GGC TGC CCA CCA   478

 R   M   F   G   V   N   C   S   Q   L   C   Q   C   D   L   G   E   M   C   H   179
AGA ATG TTT GGT GTC AAC TGC TCC CAG CTA TGT CAG TGT GAT CTC GGA GAG ATG TGC CAC   538

 P   E   T   G   A   C   V   C   P   P   G   H   S   G   A   D   C   K   M   G   199
CCA GAG ACT GGG GCT TGT GTC TGT CCC CCA GGA CAC AGT GGT GCA GAC TGC AAA ATG GGA   598

 S   Q   E   S   F   T   I   M   P   T   S   P   V   T   H   N   S   L   G   A   219
AGC CAG GAG TCC TTC ACC ATA ATG CCC ACC TCT CCC GTG ACC CAT AAC TCA CTG GGT GCA   658

 V   I   G   I   A   V   L   G   T   L   V   V   A   L   I   A   L   F   I   G   239
GTG ATT GGC ATT GCA GTA CTG GGA ACC CTC GTG GTG GCC CTG ATA GCA CTG TTC ATT GGC   718
```

FIG. 16A

```
 Y  R  Q  W  Q  K  G  K  E  H  E  H  L  A  V  A  Y  S  T  G   259
TAC CGC CAG TGG CAA AAG GGC AAG GAA CAT GAG CAC TTG GCA GTG GCT TAC AGC ACT GGG  778

 R  L  D  G  S  D  Y  V  M  P  D  V  S  P  S  Y  S  H  Y  Y   279
CGG CTG GAT GGC TCT GAT TAC GTC ATG CCA GAT GTC TCT CCG AGC TAT AGT CAC TAC TAC  838

 S  N  P  S  Y  H  T  L  S  Q  C  S  P  N  P  P  P  P  N  K   299
TCC AAC CCC AGC TAC CAC ACA CTG TCT CAG TGT TCT CCT AAC CCC CCG CCC CCT AAC AAG  898

 V  P  G  S  Q  L  F  V  S  S  Q  A  P  E  R  P  S  R  A  H   319
GTC CCA GGC AGT CAG CTC TTT GTC AGC TCT CAG GCC CCT GAG CGG CCA AGC AGA GCC CAC  958

 G  R  E  N  H  T  T  L  P  A  D  W  K  H  R  R  E  P  H  D   339
GGG CGT GAG AAC CAT ACC ACA CTG CCC GCT GAC TGG AAG CAC CGC CGG GAG CCC CAT GAC 1018

 R  G  A  S  H  L  D  R  S  Y  S  C  S  Y  S  H  R  N  G  P   359
AGA GGC GCC AGC CAC CTG GAC CGA AGC TAT AGC TGT AGC TAT AGC CAC AGG AAT GGC CCA 1078

 G  P  F  C  H  K  G  P  I  S  E  E  G  L  G  A  S  V  M  S   379
GGA CCA TTC TGT CAT AAA GGT CCC ATC TCT GAA GAG GGA CTA GGG GCA AGC GTT ATG TCC 1138

 L  S  S  E  N  P  Y  A  T  I  R  D  L  P  S  L  P  G  E  P   399
CTG AGC AGT GAG AAC CCC TAT GCT ACC ATC CGA GAC CTG CCC AGC CTG CCT GGG GAA CCC 1198

 R  E  S  G  Y  V  E  M  K  G  P  P  S  V  S  P  P  R  Q  S   419
CGA GAA AGT GGC TAT GTG GAG ATG AAA GGA CCT CCA TCA GTG TCC CCT CCC AGG CAG TCT 1258

 L  H  L  R  D  R  Q  Q  R  Q  L  Q  P  Q  R  D  S  G  T  Y   439
CTT CAT CTC CGG GAC AGG CAG CAG CGG CAA CTG CAG CCA CAG AGG GAC AGC GGC ACC TAT 1318

 E  Q  P  S  P  L  S  H  N  E  E  S  L  G  S  T  P  P  L  P   459
GAG CAG CCC AGC CCC TTG AGC CAT AAT GAA GAG TCT TTG GGC TCC ACG CCC CCG CTT CCT 1378

 P  G  L  P  P  G  H  Y  D  S  P  K  N  S  H  I  P  G  H  Y   479
CCA GGC CTG CCT CCT GGT CAC TAC GAC TCC CCC AAG AAC AGC CAT ATC CCT GGA CAC TAT 1438

 D  L  P  P  V  R  H  P  P  S  P  P  S  R  R  Q  D  R  *      498
GAC TTG CCT CCA GTA CGG CAT CCT CCA TCC CCT CCA TCC CGG CGC CAG GAC CGC TGA     1495
```

FIG.16B

AGAGCCGGCATGGTATGGGAGCGTGCCTATGTACCTTGCCAGGAGCAGGGACTGGACCAGCAGGCCACGAACAGAAACA 1574

CTTGGTGAAGTGAACAGAGACGGACTGTGGCCCTGTGCTTCCACCGAGGGAGACACTAGTTGACAAAGTGTCTAACCCT 1653

CTTTTCCAACCCACTGCTCAAGTCCCTGTGGACATAAGCTGGTGGGCAGAATGTTGTTGTACAAGTGTGATTTTAGATC 1732

GATTTTTTTTTAAAGTATGTGTTGGGTACCTTTTCTGTGTGTATGCTCAGGCAGGCTGTGTGTGTCTCTAGTTGGCTTT 1811

AGAGGGAGTCAGGTATAGGTTCTGCCTTCTGCACTTTCCATCTTATCTAGTAGTCAGCTTCCAAGCTTAACTAGTTAGA 1890

GCTCCACCAGCAGCAGGCCCTAACTACCTGCCTGCCCTTCACCCAGTAATCCTCCATGTCTTTGCTCAGAGGATTGCTC 1969

CCCGACTCTGGTGTTGTCCTCCTGGTACGCCTTGACGGTCCTGCAGTCTCCCTTTCCCGTCTTGCTTCATTCTTTCCCA 2048

GAATGAAGGCTGTCTGCCACCCTACTTCCCAGCCCAGGAATTGGCACATCTAAGTTCAGCCTTCCTAAGTTACCCGTTG 2127

AGTCCTGCTTGCCCTTCACATATTCCACAGAACACCCACCCCACATCTGCTTCATAGCTACTCTCTTCTCCACGTACCC 2206

ACAGAAGGCAGAAGTGGTACCAGGCAAGAAGATGGGATTGTTGCATTTTGTTTTGTTTTTGAGACTCTGTCTCACTATG 2285

TAGTCCTGGCTGGCCTGGAACTCAAGAGCTCTGCCTGCCTCTGCCTCTTGAGTGCTGGGTTTAACGGCTCAGGGTCACA 2364

TGCACAGCTCAAGCTGCACTCCGATGTGCTTTCCCCTGTTGCTAGATTAGCGTCTGCCTCCCCCTAGTGGAGAGGCTGA 2443

TCGCCAGCTCTCTGATGCAGGACTCTGGTGTTTAGGCTCACTCACTATTGGTTTCCTTGGCACAGGGTAGTCACTCAAT 2522

AAATGTTCCTCTAAAAGCTGAAAAAAAAAAAAAAAAAGGGCGGCCGC 2569

FIG.16C

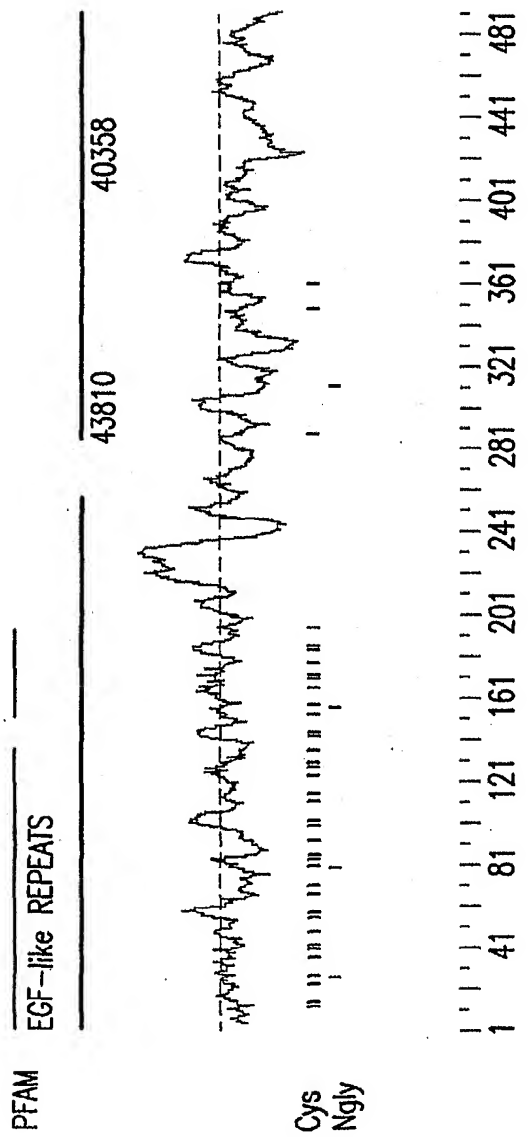


FIG.17

Input file T295Athyb23d9; Output File T295Athyb23d9.pat
Sequence length 1497

```
GTCGACCCACGCGTCCGGCTCCCAGCCCACCCCCAAACAGACACAGCGTAGCCCGGGCCAGCTCTTAAGGAGTTCAGGA  79

GTGAGAAGAGGCCCTCAGAGATCTGACAGCCTAGGAGTGCGTGGACACCACCTCAGCCCACTGAGCAGGAGTCACAGCA 158

                                                                M  A  P  A  R    5
CGAAGACCAAGCGCAAAGCGACCCCTGCCCTCCATCCTGACTGCTCCTCCTAAGAGAG ATG GCA CCG GCC AGA  231

 A  G  F  C  P  L  L  L  L  L  L  L  G  L  W  V  A  E  I  P   25
GCA GGA TTC TGC CCC CTT CTG CTG CTT CTG CTG CTG GGG CTG TGG GTG GCA GAG ATC CCA  291

 V  S  A  K  P  K  G  M  T  S  S  Q  W  F  K  I  Q  H  M  Q   45
GTC AGT GCC AAG CCC AAG GGC ATG ACC TCA TCA CAG TGG TTT AAA ATT CAG CAC ATG CAG  351

 P  S  P  Q  A  C  N  S  A  M  K  N  I  N  K  H  T  K  R  C   65
CCC AGC CCT CAA GCA TGC AAC TCA GCC ATG AAA AAC ATT AAC AAG CAC ACA AAA CGG TGC  411

 K  D  L  N  T  F  L  H  E  P  F  S  S  V  A  A  T  C  Q  T   85
AAA GAC CTC AAC ACC TTC CTG CAC GAG CCT TTC TCC AGT GTG GCC GCC ACC TGC CAG ACC  471

 P  K  I  A  C  K  N  G  D  K  N  C  H  Q  S  H  G  P  V  S  105
CCC AAA ATA GCC TGC AAG AAT GGC GAT AAA AAC TGC CAC CAG AGC CAC GGG CCC GTG TCC  531

 L  T  M  C  K  L  T  S  G  K  Y  P  N  C  R  Y  K  E  K  R  125
CTG ACC ATG TGT AAG CTC ACC TCA GGG AAG TAT CCG AAC TGC AGG TAC AAA GAG AAG CGA  591

 Q  N  K  S  Y  V  V  A  C  K  P  P  Q  K  K  D  S  Q  Q  F  145
CAG AAC AAG TCT TAC GTA GTG GCC TGT AAG CCT CCC CAG AAA AAG GAC TCT CAG CAA TTC  651

 H  L  V  P  V  H  L  D  R  V  L  *                          157
CAC CTG GTT CCT GTA CAC TTG GAC AGA GTC CTT TAG                687
```

FIG.18A

GTTTCCAGACTGGCTTGCTCTTTGGCTGACCTTCAATTCCCTCTCCAGGACTCCGCACCACTCCCCTACACCCAGAGCA 766

TTCTCTTCCCCTCATCTCTTGGGGCTGTTCCTGGTTCAGCCTCTGCTGGGAGGCTGAAGCTGACACTCTGGTGAGCTGA 845

GCTCTAGAGGGATGGCTTTTCATCTTTTTGTTGCTGTTTTCCCAGATGCTTATCCCCAAGAAACAGCAAGCTCAGGTCT 924

GTGGGTTCCCTGGTCTATGCCATTGCACATGTCTCCCCTGCCCCCTGGCATTAGGGCAGCATGACAAGGAGAGGAAATA 1003

AATGGAAAGGGGGCATATGGGATTTGTGGACACAGCTGTTTCTGTTCCTGAACTAGAAGTCTTCCCCAGCTCTGACGTG 1082

GCAGTGAGGTGACCTGAAGGAAAGAAAAATATAAATAAATACCACTTCATATTTGTATAGAATCCTCTAATCCCTTGTG 1161

ACATAGACTTGACAGGGATTGTATGCCTTCTTTATGGATGAGGAAATTAAGGTTTTAGAAAGCTTAATGAATTAAAGAG 1240

CTTGTCTAATTAGTTAGTAGCAGAACCTGGACTTGAACCTAGGTCTCCTTGCTCTAAATACAGTGTACCTTCTACTCTA 1319

CCAGTTGCGCAAGAAAGAAGTCACTGTTACAGAGGCAAGCGGTGAACTAGGTAAGAGTTCACTCATGAAGAAACGAGTG 1398

CTCTGAAGAGCCAGTTACCCTGTGTTGGCTGCAATAAAGGTCATTACCTCTCTAGCCAAAAAAAAAAAAAAAAAAAAA 1477

AAAAAAAAAAAAAAAAAAA 1497

FIG.18B

T295

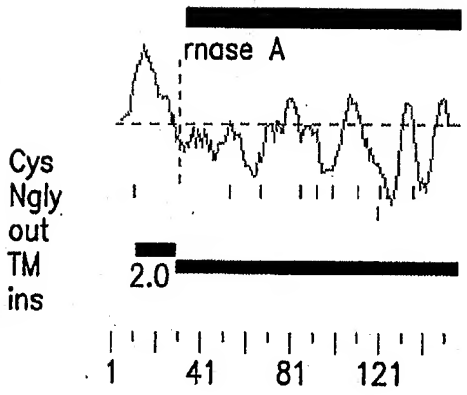


FIG.19

```
     *->qesrAqkFlrQHiDspktsssnpnYCNqMMdkrRnmtqgrCKpvNTF
        + ++ q+F++QH+   ++s +    CN +M k++n     rCK+ NTF
 32     GMTSSQWFKIQHM---QPSPQA---CNSAM-KNINKHTKRCKDLNTF   71

     vHesladVkaVCsqkNvtCkNGqkNCyqSkssfqiTdCrltggsqkyPnC
     +He++++V a C ++ + CkNG kNC+qS+ ++++T C+lt+g   yPnC
 72  LHEPFSSVAATCQTPKIACKNGDKNCHQSHGPVSLTMCKLTSGK--YPNC  119

     rYrtsastkhIiVACEgrd.rddPyynPyvPVHFDasv<-*
     rY+ + ++k ++VAC +++++d+ ++   vPVH+D++
120  RYKEKRQNKSYVVACKPPQkKDSQQFH-LVPVHLDRVL    156
```

FIG.20

Input file T354Athla42a4; Output File T354Athla42a4.pat
Sequence length 1788

```
                                                                M  P  L  L      4
GTCGACCCACGGCGTCCGGCCAGGCTCCACTGAGGGGAACGGGGACCTGTCTGAAGAGAAG ATG CCC CTG CTG     73

 T  L  Y  L  L  L  F  W  L  S  G  Y  S  I  A  T  Q  I  T  G    24
ACA CTC TAC CTG CTC CTC TTC TGG CTC TCA GGC TAC TCC ATT GCC ACT CAA ATC ACC GGT  133

 P  T  T  V  N  G  L  E  R  G  S  L  T  V  Q  C  V  Y  R  S    44
CCA ACA ACA GTG AAT GGC TTG GAG CGG GGC TCC TTG ACC GTG CAG TGT GTT TAC AGA TCA  193

 G  W  E  T  Y  L  K  W  W  C  R  G  A  I  W  R  D  C  K  I    64
GGC TGG GAG ACC TAC TTG AAG TGG TGG TGT CGA GGA GCT ATT TGG CGT GAC TGC AAG ATC  253

 L  V  K  T  S  G  S  E  Q  E  V  K  R  D  R  V  S  I  K  D    84
CTT GTT AAA ACC AGT GGG TCA GAG CAG GAG GTG AAG AGG GAC CGG GTG TCC ATC AAG GAC  313

 N  Q  K  N  R  T  F  T  V  T  M  E  D  L  M  K  T  D  A  D   104
AAT CAG AAA AAC CGC ACG TTC ACT GTG ACC ATG GAG GAT CTC ATG AAA ACT GAT GCT GAC  373

 T  Y  W  C  G  I  E  K  T  G  N  D  L  G  V  T  V  Q  V  T   124
ACT TAC TGG TGT GGA ATT GAG AAA ACT GGA AAT GAC CTT GGG GTC ACA GTT CAA GTG ACC  433

 I  D  P  A  S  T  P  A  P  T  T  P  T  S  T  T  F  T  A  P   144
ATT GAC CCA GCG TCG ACT CCT GCC CCC ACC ACG CCT ACT TCC ACT ACG TTT ACA GCA CCA  493

 V  T  Q  E  E  T  S  S  S  P  T  L  T  G  H  H  L  D  N  R   164
GTC ACC CAA GAA GAA ACT AGC AGC TCC CCA ACT CTG ACC GGC CAC CAC TTG GAC AAC AGG  553

 H  K  L  L  K  L  S  V  L  L  P  L  I  F  T  I  L  L  L  L   184
CAC AAG CTC CTG AAG CTC AGT GTC CTC CTG CCC CTC ATC TTC ACC ATA TTG CTG CTG CTT  613

 L  V  A  A  S  L  L  A  W  R  M  M  K  Y  Q  Q  K  A  A  G   204
TTG GTG GCC GCC TCA CTC TTG GCT TGG AGG ATG ATG AAG TAC CAG CAG AAA GCA GCC GGG  673

 M  S  P  E  Q  V  L  Q  P  L  E  G  D  L  C  Y  A  D  L  T   224
ATG TCC CCA GAG CAG GTA CTG CAG CCC CTG GAG GGC GAC CTC TGC TAT GCA GAC CTG ACC  733
```

FIG.21A

```
 L  Q  L  A  G  T  S  P  R  K  A  T  T  K  L  S  S  A  Q  V  244
CTG CAG CTG GCC GGA ACC TCC CCG CGA AAG GCT ACC ACG AAG CTT TCC TCT GCC CAG GTT  793

 D  Q  V  E  V  E  Y  V  T  M  A  S  L  P  K  E  D  I  S  Y  264
GAC CAG GTG GAA GTG GAA TAT GTC ACC ATG GCT TCC TTG CCG AAG GAG GAC ATT TCC TAT  853

 A  S  L  T  L  G  A  E  D  Q  E  P  T  Y  C  N  M  G  H  L  284
GCA TCT CTG ACC TTG GGT GCT GAG GAT CAG GAA CCG ACC TAC TGC AAC ATG GGC CAC CTC  913

 S  S  H  L  P  G  R  G  P  E  E  P  T  E  Y  S  T  I  S  R  304
AGT AGC CAC CTC CCC GGC AGG GGC CCT GAG GAG CCC ACG GAA TAC AGC ACC ATC AGC AGG  973

 P  *                                                          306
CCT TAG                                                       979

CCTGCACTCCAGGCTCCTTCTTGGACCCCAGGCTGTGAGCACACTCCTGCCTCATCGACCGTCTGCCCCCTGCTCCCCT 1058

CATCAGGACCAACCCCGGGGACTGGTGCCTCTGCCTGATCAGCCAGCATTGCCCCTAGCTCTGGGTTGGGCTTGGGGCCA 1137

AGTCTCAGGGGGCTTCTAGGAGTTGGGGTTTTCTAAACGTCCCCTCCTCTCCTACATAGTTGAGGAGGGGGCTAGGGAT 1216

ATGCTCTGGGGCTTTCATGGGAATGATGAAGATGATAATGAGAAAAATGTTATCATTATTATCATGAAGTACCATTATC 1295

ATAATACAATGAACCTTTATTTATTGCCTACCACATGTTATGGGCTGAATAATGGCCCCCAAAGATATCTGTGTCCTAA 1374

TCCTCAGAACTTGTGACTGTTACCTTCTGTGGCAGAAAGGGACAGTGCAGATGTATGTAAGTTAAGGACTTTGAGATAG 1453

AGAGGTTATTCTTGCTGATTCAGGTGGGCCCAAAATATCACCACAAGGGTCCTCATAAGAAAGAGGCCAGAAGGTCAAA 1532

GAGGTAGAGACAAAGTGATGATGGAAGTGGACGTGGGTGTGACGTGAGCAGGGGCCATGAATGCCGCAGCCTTCAGATG 1611

CCAGAAAGGGAAAGGAATGGATTCCCCTGCCTGGAGCCTCCAAAAGAAACCAGCCCTGCCCACGCCTTGACTTGAGCCC 1690

ATTGAAACTGATCTTGAGCTCCTGGCCTCCAGAATTGCAGGAGAATAAATTTGTGTTGTTTTTAAAAAAAAAAAAAAAA 1769

AAAAAGGGCGGCCGCTAGA                                            1788
```

FIG.21B

T354

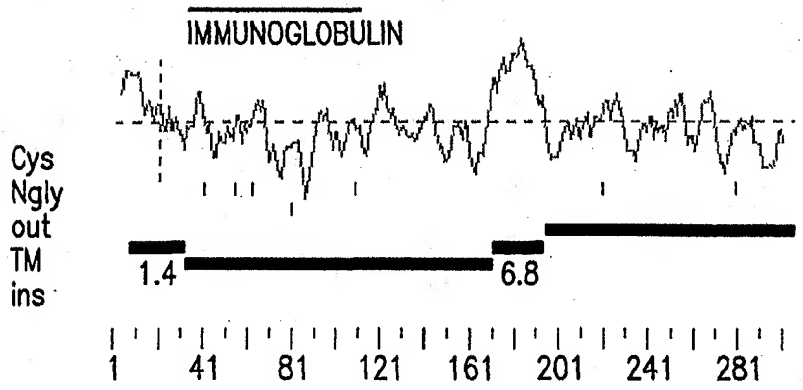

IMMUNOGLOBULIN
Cys
Ngly
out
TM
ins
1.4
6.8
1
41
81
121
161
201
241
281

FIG.22

```
   *->GesvtLtCsvsgfgppgvsvtWyf.....kngk.lgpsllgysysrl
      ++s+t +C ++  + + +++ W+ ++   ++ k l    ++ s   +
33    RGSLTVQCVYR--SGWETYLKWWCrgaiwRDCKiLVK--TSGSEQEV 75

   esgekanlsegrfsis.....sltLtissvekeDsGtYtCvv<-*
    ++        r+si ++++++++t+t+ ++ k D+ tY+C
76 KRD--------RVSIKdnqknrTFTVTMEDLMKTDADTYWCGI    110
```

FIG.23

Input file T378Athta28f4; Output File T378Athta28f4.pat
Sequence length 3258

```
                                                  M  D  H  C  G  A  L  F  L      9
CACGCGTCCGGCCAGTTCTTGGAGGAGACTCTGCACAGGGC ATG GAT CAC TGT GGT GCC CTT TTC CTG    68

 C  L  C  L  L  T  L  Q  N  A  T  T  E  T  W  E  E  L  L  S     29
TGC CTG TGC CTT CTG ACT TTG CAG AAT GCA ACA ACA GAG ACA TGG GAA GAA CTC CTG AGC   128

 Y  M  E  N  M  Q  V  S  R  G  R  S  S  V  F  S  S  R  Q  L     49
TAC ATG GAG AAT ATG CAG GTG TCC AGG GGC CGG AGC TCA GTT TTT TCC TCT CGT CAA CTC   188

 H  Q  L  E  Q  M  L  L  N  T  S  F  P  G  Y  N  L  T  L  Q     69
CAC CAG CTG GAG CAG ATG CTA CTG AAC ACC AGC TTC CCA GGC TAC AAC CTG ACC TTG CAG   248

 T  P  T  I  Q  S  L  A  F  K  L  S  C  D  F  S  G  L  S  L     89
ACA CCC ACC ATC CAG TCT CTG GCC TTC AAG CTG AGC TGT GAC TTC TCT GGC CTC TCG CTG   308

 T  S  A  T  L  K  R  V  P  Q  A  G  G  Q  H  A  R  G  Q  H    109
ACC AGT GCC ACT CTG AAG CGG GTG CCC CAG GCA GGA GGT CAG CAT GCC CGG GGT CAG CAC   368

 A  M  Q  F  P  A  E  L  T  R  D  A  C  K  T  R  P  R  E  L    129
GCC ATG CAG TTC CCC GCC GAG CTG ACC CGG GAC GCC TGC AAG ACC CGC CCC AGG GAG CTG   428

 R  L  I  C  I  Y  F  S  N  T  H  F  F  K  D  E  N  N  S  S    149
CGG CTC ATC TGT ATC TAC TTC TCC AAC ACC CAC TTT TTC AAG GAT GAA AAC AAC TCA TCT   488

 L  L  N  N  Y  V  L  G  A  Q  L  S  H  G  H  V  N  N  L  R    169
CTG CTG AAT AAC TAC GTC CTG GGG GCC CAG CTG AGT CAT GGG CAC GTG AAC AAC CTC AGG   548

 D  P  V  N  I  S  F  W  H  N  Q  S  L  E  G  Y  T  L  T  C    189
GAT CCT GTG AAC ATC AGC TTC TGG CAC AAC CAA AGC CTG GAA GGC TAC ACC CTG ACC TGT   608

 V  F  W  K  E  G  A  R  K  Q  P  W  G  G  W  S  P  E  G  C    209
GTC TTC TGG AAG GAG GGA GCC AGG AAA CAG CCC TGG GGG GGC TGG AGC CCT GAG GGC TGT   668

 R  T  E  Q  P  S  H  S  Q  V  L  C  R  C  N  H  L  T  Y  F    229
CGT ACA GAG CAG CCC TCC CAC TCT CAG GTG CTC TGC CGC TGC AAC CAC CTC ACC TAC TTT   728
```

FIG.24A

```
 A  V  L  M  Q  L  S  P  A  L  V  P  A  E  L  L  A  P  L  T   249
GCT GTT CTC ATG CAA CTC TCC CCA GCC CTG GTC CCT GCA GAG TTG CTG GCA CCT CTT ACG   788

 Y  I  S  L  V  G  C  S  I  S  I  V  A  S  L  I  T  V  L  L   269
TAC ATC TCC CTC GTG GGC TGC AGC ATC TCC ATC GTG GCC TCG CTG ATC ACA GTC CTG CTG   848

 H  F  H  F  R  K  Q  S  D  S  L  T  R  I  H  M  N  L  H  A   289
CAC TTC CAT TTC AGG AAG CAG AGT GAC TCC TTA ACA CGC ATC CAC ATG AAC CTG CAT GCC   908

 S  V  L  L  L  N  I  A  F  L  L  S  P  A  F  A  M  S  P  V   309
TCC GTG CTG CTC CTG AAC ATC GCC TTC CTG CTG AGC CCC GCA TTC GCA ATG TCT CCT GTG   968

 P  G  S  A  C  T  A  L  A  A  A  L  H  Y  A  L  L  S  C  L   329
CCC GGG TCA GCA TGC ACG GCT CTG GCC GCT GCC CTG CAC TAC GCG CTG CTC AGC TGC CTC  1028

 T  W  M  A  I  E  G  F  N  L  Y  L  L  L  G  R  V  Y  N  I   349
ACC TGG ATG GCC ATC GAG GGC TTC AAC CTC TAC CTC CTC CTC GGG CGT GTC TAC AAC ATC  1088

 Y  I  R  R  Y  V  F  K  L  G  V  L  G  W  G  A  P  A  L  L   369
TAC ATC CGC AGA TAT GTG TTC AAG CTT GGT GTG CTA GGC TGG GGG GCC CCA GCC CTC CTG  1148

 V  L  L  S  L  S  V  K  S  S  V  Y  G  P  C  T  I  P  V  F   389
GTG CTG CTT TCC CTC TCT GTC AAG AGC TCG GTA TAC GGA CCC TGC ACA ATC CCC GTC TTC  1208

 D  S  W  E  N  G  T  G  F  Q  N  M  S  I  C  W  V  R  S  P   409
GAC AGC TGG GAG AAT GGC ACA GGC TTC CAG AAC ATG TCC ATA TGC TGG GTG CGG AGC CCC  1268

 V  V  H  S  V  L  V  M  G  Y  G  G  L  T  S  L  F  N  L  V   429
GTG GTG CAC AGT GTC CTG GTC ATG GGC TAC GGC GGC CTC ACG TCC CTC TTC AAC CTG GTG  1328

 V  L  A  W  A  L  W  T  L  R  R  L  R  E  R  A  D  A  P  S   449
GTG CTG GCC TGG GCG CTG TGG ACC CTG CGC AGG CTG CGG GAG CGG GCG GAT GCA CCA AGT  1388

 V  R  A  C  H  D  T  V  T  V  L  G  L  T  V  L  L  G  T  T   469
GTC AGG GCC TGC CAT GAC ACT GTC ACT GTG CTG GGC CTC ACC GTG CTG CTG GGA ACC ACC  1448

 W  A  L  A  F  F  S  F  G  V  F  L  L  P  Q  L  F  L  F  T   489
TGG GCC TTG GCC TTC TTT TCT TTT GGC GTC TTC CTG CTG CCC CAG CTG TTC CTC TTC ACC  1508
```

FIG.24B

```
 I  L  N  S  L  Y  G  F  F  L  F  L  W  F  C  S  Q  R  C  R   509
ATC TTA AAC TCG CTC TAC GGT TTC TTC CTT TTC CTG TGG TTC TGC TCC CAG CGG TGC CGC  1568

 S  E  A  E  A  K  A  Q  I  E  A  F  S  S  S  Q  T  T  Q  *   529
TCA GAA GCA GAG GCC AAG GCA CAG ATA GAG GCC TTC AGC TCC TCC CAA ACA ACA CAG TAG  1628

TCCGGGCCTCCTGGCCTGGAATCCTCAGCCTCTCTGGCCGCCAGTAGCCTGAGGCTACGGCTCCTGCTAGAGAGGGTGG  1707

CAGGCCTGCTGCTGGACCCCAGAGGCCACTGTGACCGCCAAGGGGCCTTTTCCACTTCCACGGCCTCTCCAGGCACTGA  1786

GGGGAAGGCATTGCTCTACCTCTCCCTGACATTTTGCTCCGGGGCAGATCCAACCTTACCTGGGGCAGCAAACTTTGTC  1865

CTGGTACCTGGGCCCAGCTCGCCAGGGATGTGGGCAGAGCACCAGCCTGGGCATCAGGAAGCCAAGTTTCAAGGACTGT  1944

CTTTGAGTCTGTCTGTATGACCTTGGGCCTGCCACTTCTCACAGACCCTAGGTATCCACAGCTGTGACATGGGGGCAAG  2023

CGGCTTTGTTTCAGCCTAACCCAGGAGCTTAGTAAAAATTGCATAAGACCAGGGGGAAGAGTGTCAGCGTGGGGTGGGA  2102

ATTCCCGCGGCCTCCACCTGCTTGCTAGGGGCAGGATCTCATTCAGGCTGCCCTGGAAGCACCTGCTTGGCCCTGCCAC  2181

CTTCCTCCAGGGGAGGGCCAGATGGCATCCTGGCTTGGGGCGGGTGGGACCTACCCAGGCTCTGAGACTTTACTGGCCT  2260

ATGCCTGAGGCCTCTTTTCCTTTAACTCCCTAAATTATGATGACTCCAAGTCCAAGCCCACCCTTCCCAAAGATTGGGA  2339

GGTTCCGCCGTTCCCAGAGGCTCCTCCTGCGGTGCTCCCAAGACTTCCATAGACCATCTGGACCAGTAGCCCATCCCGC  2418

AGTTTTCTTGGGGGCAGAGGAAAACGCTTCTTTCTCCTCCAGCTGAATCAGCTGGATCCCAGTGTCCTGGCTGTTTGGT  2497

GATTGGGCAAGATTGAATTTGCCCAGGTAGGCGTGAGAGTGTGGGTTTTAAATTCGAAGCTCAGGCCATAGTTTCAGAG  2576

AATCACCCTTACCCCAGACCTTCATGAGACAGTGCTCATGAAGCCAGTGCGTTTCCCAGAACGAACACTAGGCGGCACC  2655

GTTGGTCCACACTCAGAGGCCCTTGGCGCCAAGACTGCATCTAGAATCGCTCAAACACCTGTTTGCAGACCCCATGCAC  2734

CAGCTGGAGGGGCCGTAACTGCAGGACTGCGCCTACTGAGTGACCCATTTCCTCCAGGAGGAAAGGCAAGACACGCTTA  2813

CACGGCCATTTGTCTCTTTTCCCAATGCGGCGGTGCACTTTCGCTCTTGGGGGCTGCACCCCAGACATAGCTGGCACCA  2892
```

FIG.24C

GAGCAGGGTGCTCAGGTGGTGGGTGCTCAGGGCCCTGCCCCAGGCCACTGGGCCGTTTTGATGACCTCGAAGGTCACAG 2971

GCAGAAAATAGGAGCAGGATTTCCCCTGGGGAAAAGTTCTCCTGGGACATCTTCTGCTCTTCTGTACATTTCTAGATGC 3050

AAATAACTCCTTCACCAGGCAGTGAGTGGCGTAGGCTCTGGAGCCAGGCTGCCTGGGCTCCAATGCCAGCTCTGCCACT 3129

TGCTAGCTGTGAGACTGTGGACAAACCACTCAGCCTCTGTGTGCCTCAGTTTTCCTATTTGTAAAATAGAGGCCATAGT 3208

GGTACCTATTTTGAAGACTAAGTAAAAGAATTCAAATAAAGAGACTTGGC 3258

FIG.24D

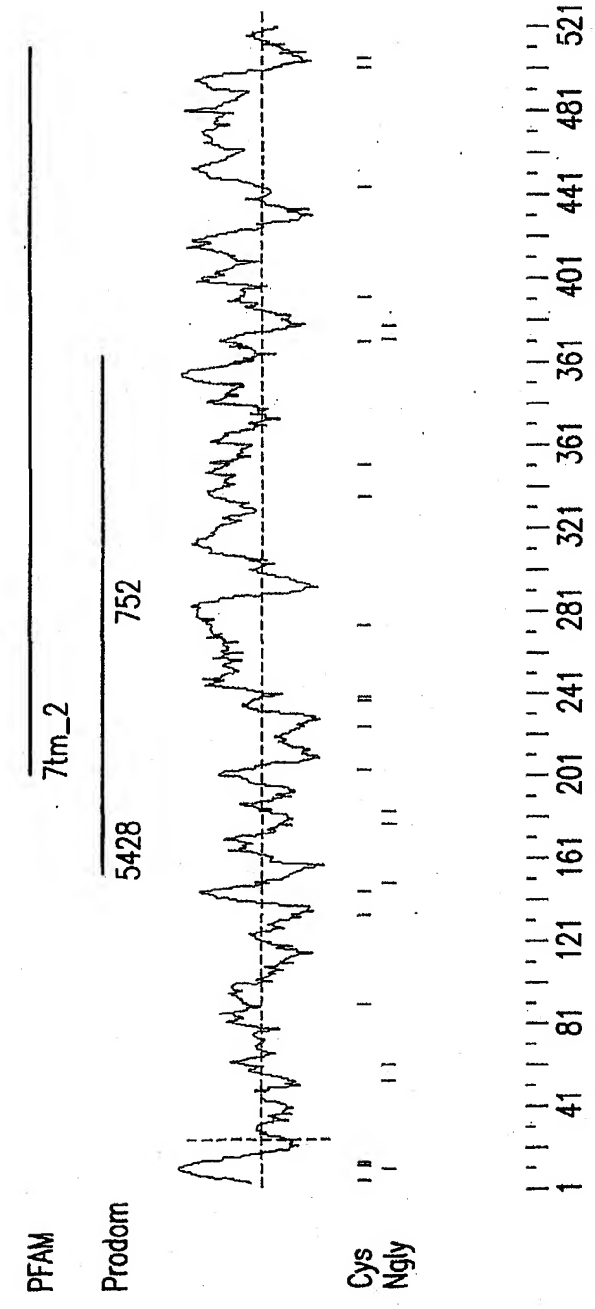


FIG.25

```
    *->CnrtWDgitC..Wpdt....ppGelVvvpCPkyfygfssdqtdttgn
              +tC  W+ + +++p+G ++   C       + +q   + +
187    -------LTCvfWKEGarkqPWGGWSPEGC-------RTEQ---PSH 216

    vsRnCtedGsWsepppsNrtWrnysaCgeddpeeesekkkkyylvlkiiY
    ++  C+ +    +++              + ++ +    +++ +  l +i
217 SQVLCRCNH--LTYFA-----------VLMQLSPALVPAELLAPLTYIS 252

    tvGYSlSLaaLlvAvvILllFRkLhtlwpdnadgalevgapWGAPfqvrr

    +vG S+S++a l+ v++   FRk       + +
253 LVGCSISIVASLITVLLHFHFRKQS-----DSL---------------- 280

    SirCtRNyIHmNLFlSFILrAasvfikdavlksevssdeperLssrcsls
        tR  IHmNL +S +L +++ ++  a   s v+ ++
281 ----TR--IHMNLHASVLLLNIAFLLSPAFAMSPVPGSA---------- 313

    tgqvvvgCkllvvfQfqYcvmtNffWllvEGlYLhtLLvvtffsErkylw
          C +l ++ ++Y++++  +W+ +EG  L+ LL   +   ++y +
314 -------CTALAAA-LHYALLSCLTWMAIEGFNLYLLLGRVY---NIYIR 352

    wYl....lIGWGvPlVfvtvWaivRllfedtgCWdsnGLAmFPEAKmCiW
     Y+ + +++GWG+P++ v     v++ ++  +C++++    F
353 RYVfklgVLGWGAPALLVLLSLSVKSSVY-GPCTIPV----FDSWENGTG 397

    msdnshlwWIIkgPiLlsilV.........NFflFinIirILvtKLRaa
       n+++ W+ + P++ s+lV + ++ ++  N++++   ++ L + LR+
398 F-QNMSICWV-RSPVVHSVLVmgyggltslfNLVVLAWALWTL-RRLRER 444

    qtgetdqrqYsqYrkLaKSTLllPLfGIhyvvFafrPsndarGvlrkik
      + +        +  +   L L  L+G++  + +f+++     v+   +
445 ADAPSVR-----ACHDTVTVLGLTVLLGTTWALAFFSFG-----VFLLPQ 484

    lyfelsLgSFQGFfVAvlYCFlNgEVQaEirrrW<-*
    l++   L+S+ GFf  ++  F+ +  ++E   +
485 LFLFTILNSLYGFF--LFLWFCSQRCRSEAEAKA     516
```

FIG.26

```
                10        20        30        40        50        60        70
inputs ATGACGCCGAGCCCCCTGTTGCTGCTCCTGCTGCCGCCGCTGCTGCTGGGGGCCTTCCCGCCGGCCGCCG
       ------------------------------------------------------------------------

                80        90       100       110       120       130       140
inputs CCGCCCGAGGCCCCCCAAAGATGGCGGACAAGGTGGTCCCACGGCAGGTGGCCCGGCTGGGCCGCACTGT
                                                 ::::
       ----------------------------------------CACG------------------------

               150       160       170       180       190       200       210
inputs GCGGCTGCAGTGCCCAGTGGAGGGGGACCCGCCGCCGCTGACCATGTGGACCAAGGATGGCCGCACCATC
       ------------------------------------------------------------------------

               220       230       240       250       260       270       280
inputs CACAGCGGCTGGAGCCGCTTCCGCGTGCTGCCGCAGGGGCTGAAGGTGAAGCAGGTGGAGCGGGAGGATG
       ------------------------------------------------------------------------

               290       300       310       320       330       340       350
inputs CCGGCGTGTACGTGTGCAAGGCCACCAACGGCTTCGGCAGCCTGAGCGTCAACTACACCCTCGTCGTGCT
       ------------------------------------------------------------------------

               360       370       380       390       400       410       420
inputs GGATGACATTAGCCCAGGGAAGGAGAGCCTGGGGCCCGACAGCTCCTCTGGGGGTCAAGAGGACCCCGCC
       ------------------------------------------------------------------------

               430       440       450       460       470       480       490
inputs AGCCAGCAGTGGGCACGACCGCGCTTCACACAGCCCTCCAAGATGAGGCGCCGGGTGATCGCACGGCCCG
                          ::::::
       ----------CGTCCG-----------------------------------------------------
                 10

               500       510       520       530       540       550       560
inputs TGGGTAGCTCCGTGCGGCTCAAGTGCGTGGCCAGCGGGCACCCTCGGCCCGACATCACGTGGATGAAGGA
       ------------------------------------------------------------------------

               570       580       590       600       610       620       630
inputs CGACCAGGCCTTGACGCGCCCAGAGGCCGCTGAGCCCAGGAAGAAGAAGTGGACACTGAGCCTGAAGAAC
       ------------------------------------------------------------------------

               640       650       660       670       680       690       700
inputs CTGCGGCCGGAGGACAGCGGCAAATACACCTGCCGCGTGTCGAACCGCGCGGGCGCCATCAACGCCACCT
```

FIG.27A

```
                710        720        730        740        750        760        770
inputs ACAAGGTGGATGTGATCCAGCGGACCCGTTCCAAGCCCGTGCTCACAGGCACGCACCCCGTGAACACGAC
       ------------------------------------------------------------------------

                780        790        800        810        820        830        840
inputs GGTGGACTTCGGGGGGACCACGTCCTTCCAGTGCAAGGTGCGCAGCGACGTGAAGCCGGTGATCCAGTGG
       ------------------------------------------------------------------------

                850        860        870        880        890        900        910
inputs CTGAAGCGCGTGGAGTACGGCGCCGAGGGCCGCCACAACTCCACCATCGATGTGGGCGGCCAGAAGTTTG
       ------------------------------------------------------------------------

                920        930        940        950        960        970        980
inputs TGGTGCTGCCCACGGGTGACGTGTGGTCGCGGCCCGACGGCTCCTACCTCAATAAGCTGCTCATCACCCG
       .......:::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
       -------GCCCACGGGTGATGTGTGGTCACGGCCTGATGGCTCCTACCTCAACAAGCTGCTCATCTCTCG
                     20         30         40         50         60         70

                990       1000       1010       1020       1030       1040       1050
inputs TGCCCGCCAGGACGATGCGGGCATGTACATCTGCCTTGGCGCCAACACCATGGGCTACAGCTTCCGCAGC
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
       GGCCCGCCAGGATGATGCTGGCATGTACATCTGCCTAGGTGCAAATACCATGGGCTACAGTTTCCGTAGC
            80         90        100        110        120        130        140

               1060       1070       1080       1090       1100       1110       1120
inputs GCCTTCCTCACCGTGCTGCCAGACCCAAAACCGCCAGGGCCACCTGTGGCCTCCTCGTCCTCGGCCACTA
       ::::::::::: :: :::::::::: ::::: :::::::::::::::::: :: :: :: :::::::
       GCCTTCCTCACTGTATTACCAGACCCCAAACCTCCAGGGCCTCCTATGGCTTCTTCATCGTCATCCACAA
          150        160        170        180        190        200        210

               1130       1140       1150       1160       1170       1180       1190
inputs GCCTGCCGTGGCCCGTGGTCATCGGCATCCCAGCCGGCGCTGTCTTCATCCTGGGCACCCTGCTCCTGTG
       ::::::: ::::: :::::: :::::::::::::: :: :::::::::::::::::: :::: :: ::
       GCCTGCCATGGCCTGTGGTGATCGGCATCCCAGCTGGTGCTGTCTTCATCCTAGGCACTGTGCTGCTCTG
          220        230        240        250        260        270        280

               1200       1210       1220       1230       1240       1250       1260
inputs GCTTTGCCAGGCCCAGAAGAAGCCGTGCACCCCCGCGCCTGCCCCTCCCCTGCCTGGGCACCGCCCGCCG
       :::::::::::: :::::::::: :::: :: :::: : ::: :::::::::: :: :: :::
       GCTTTGCCAGACCAAGAAGAAGCCATGTGCCCCAGCATCTACACTTCCTGTGCCTGGGCATCGTCCCCCA
          290        300        310        320        330        340        350

               1270       1280       1290       1300       1310       1320       1330
inputs GGGACGGCCCGCGACCGCAGCGGAGACAAGGACCTTCCCTCGTTGGCCGCCCTCAGCGCTGGCCCTGGTG
       ::::: :::: :: ::::::::::::::::::::::::::::                        ::
       GGGACATCCCGAGAACGCAGTGGTGACAAGGACCTGCCCTCATTGGC--------------------TG
          360        370        380        390        400

               1340       1350       1360       1370       1380       1390       1400
inputs TGGGGCTGTGTGAGGAGCATGGGTCTCCGGCAGCCCCCCAGCACTTACTGGGCCCAGGCCCAGTTGCTGG
       :::: :::::::::::::::: :::::::::::::::::::: :::: :: :::::::::::::::
       TGGGCATATGTGAGGAGCATGGATCCGCCATGGCCCCCCAGCACATCCTGGCCTCTGGCTCAACTGCTGG
          410        420        430        440        450        460        470
```

FIG.27B

```
          1410       1420       1430       1440       1450       1460
inputs CCCTAAGTTGTACCCCAAACTCTACACAGACATCCACACACACACACA--CACACAC--TCTCACACACA
       ::: ::: :::::::::::: ::::::: .: ::::::::::::::::::  ::::::::  :.:.::::.
       CCCCAAGCTGTACCCCAAGCTATACACAGATGTGCACACACACACACATACACACACCTGCACTCACACG
            480        490        500        510        520        530        540

       1470       1480            1490       1500       1510
inputs CTCACACGT-GGAGGGCAAGGT-C-----CACCAGCACATCCACTATCAGTGC----------------
       :::::: :: :::::::::::: : ::::::::.:::::::::::::::::
       CTCTCATGTTGGAGGGCAAGGTTCATCAACACCAGCATGTCCACTATCAGTGCTAAATACAGCGAATCTC
            550        560        570        580        590        600        610

inputs -------------

       CAAGCACTGTGTCC
```

FIG.27C

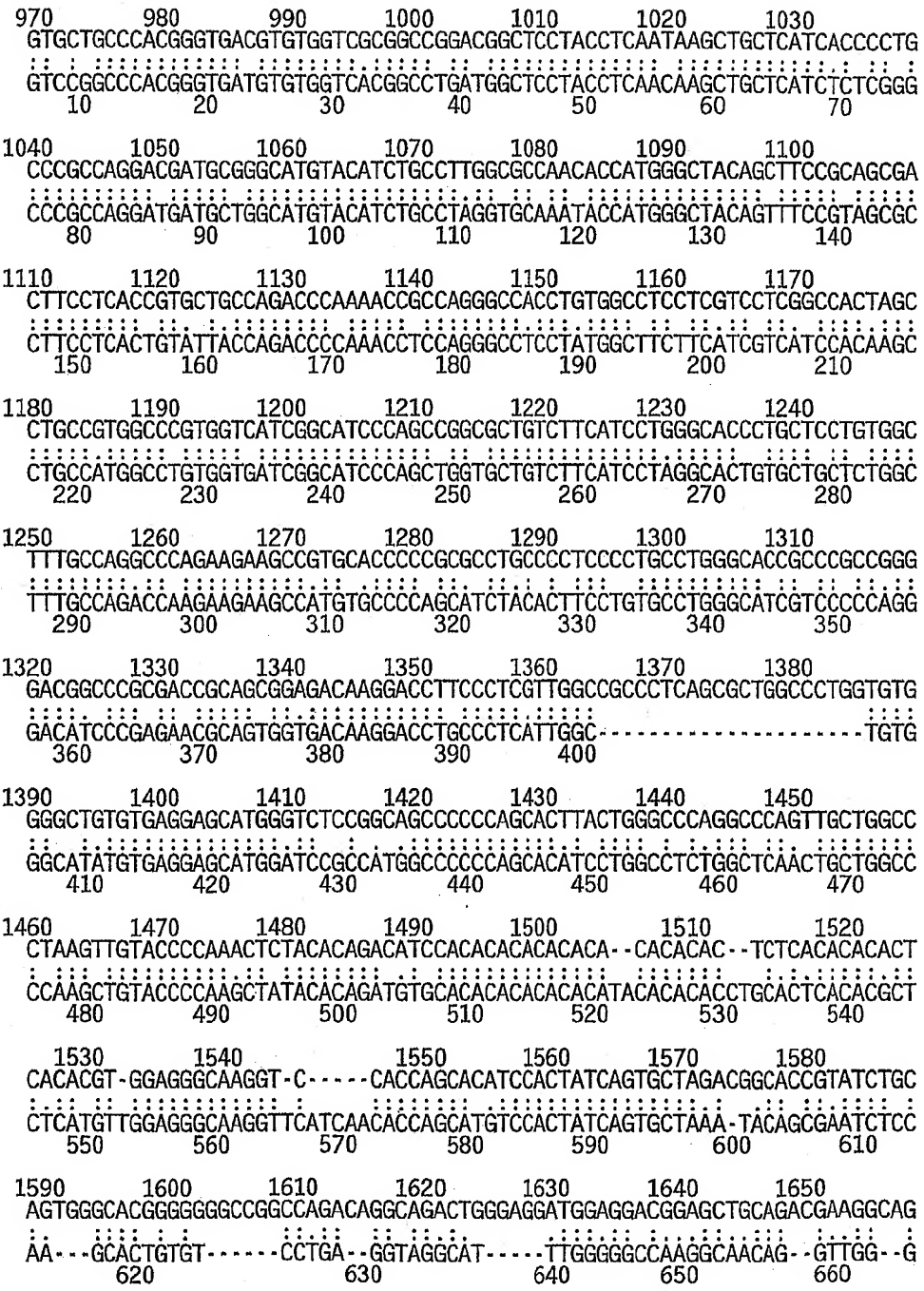


FIG.28A

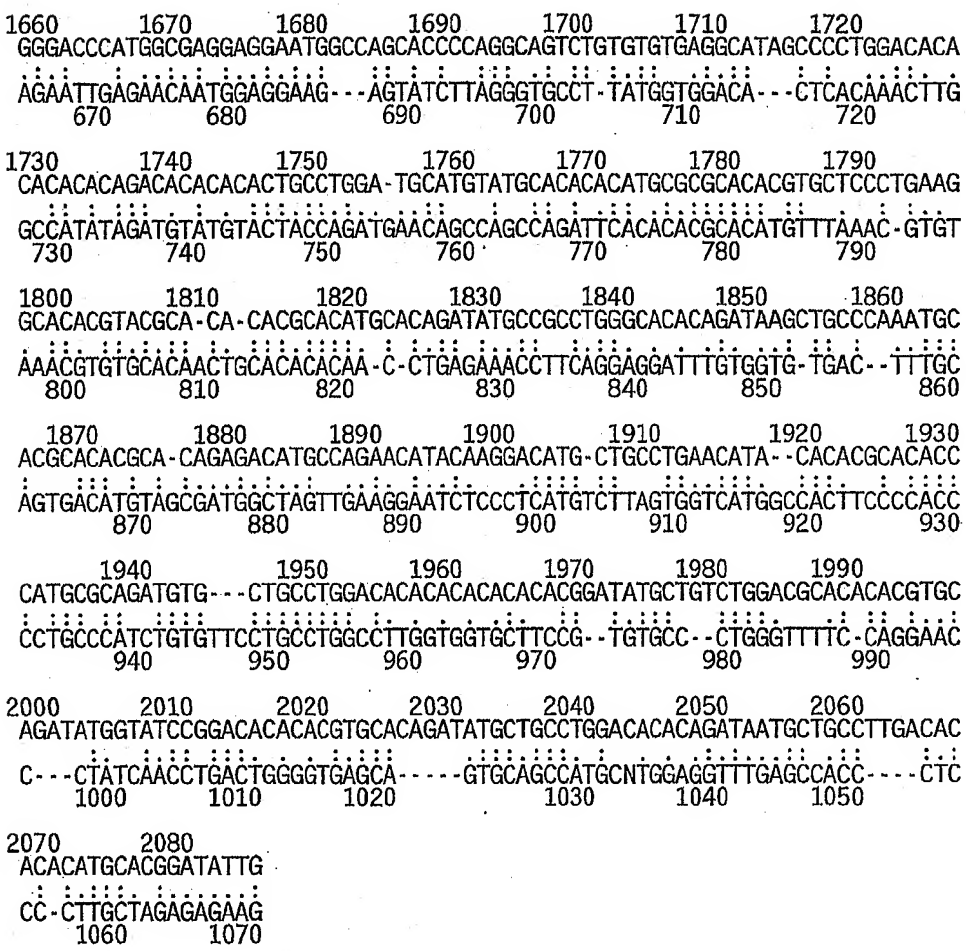

1660 1670 1680 1690 1700 1710 1720
GGGACCCATGGCGAGGAGGAATGGCCAGCACCCCAGGCAGTCTGTGTGTGAGGCATAGCCCCTGGACACA
AGAATTGAGAACAATGGAGGAAG---AGTATCTTAGGGTGCCT-TATGGTGGACA---CTCACAAACTTG
670 680 690 700 710 720
1730 1740 1750 1760 1770 1780 1790
CACACACAGACACACACACTGCCTGGA-TGCATGTATGCACACACATGCGCGCACACGTGCTCCCTGAAG
GCCATATAGATGTATGTACTACCAGATGAACAGCCAGCCAGATTCACACACGCACATGTTTAAAC-GTGT
730 740 750 760 770 780 790
1800 1810 1820 1830 1840 1850 1860
GCACACGTACGCA-CA-CACGCACATGCACAGATATGCCGCCTGGGCACACAGATAAGCTGCCCAAATGC
AAACGTGTGCACAACTGCACACACAA-C-CTGAGAAACCTTCAGGAGGATTTGTGGTG-TGAC--TTTGC
800 810 820 830 840 850 860
1870 1880 1890 1900 1910 1920 1930
ACGCACACGCA-CAGAGACATGCCAGAACATACAAGGACATG-CTGCCTGAACATA--CACACGCACACC
AGTGACATGTAGCGATGGCTAGTTGAAGGAATCTCCCTCATGTCTTAGTGGTCATGGCCACTTCCCCACC
870 880 890 900 910 920 930
1940 1950 1960 1970 1980 1990
CATGCGCAGATGTG---CTGCCTGGACACACACACACACACGGATATGCTGTCTGGACGCACACACGTGC
CCTGCCCATCTGTGTTCCTGCCTGGCCTTGGTGGTGCTTCCG--TGTGCC--CTGGGTTTTC-CAGGAAC
940 950 960 970 980 990
2000 2010 2020 2030 2040 2050 2060
AGATATGGTATCCGGACACACACGTGCACAGATATGCTGCCTGGACACACAGATAATGCTGCCTTGACAC
C---CTATCAACCTGACTGGGGTGAGCA-----GTGCAGCCATGCNTGGAGGTTTGAGCCACC----CTC
1000 1010 1020 1030 1040 1050
2070 2080
ACACATGCACGGATATTG
CC-CTTGCTAGAGAGAAG
1060 1070

FIG.28B

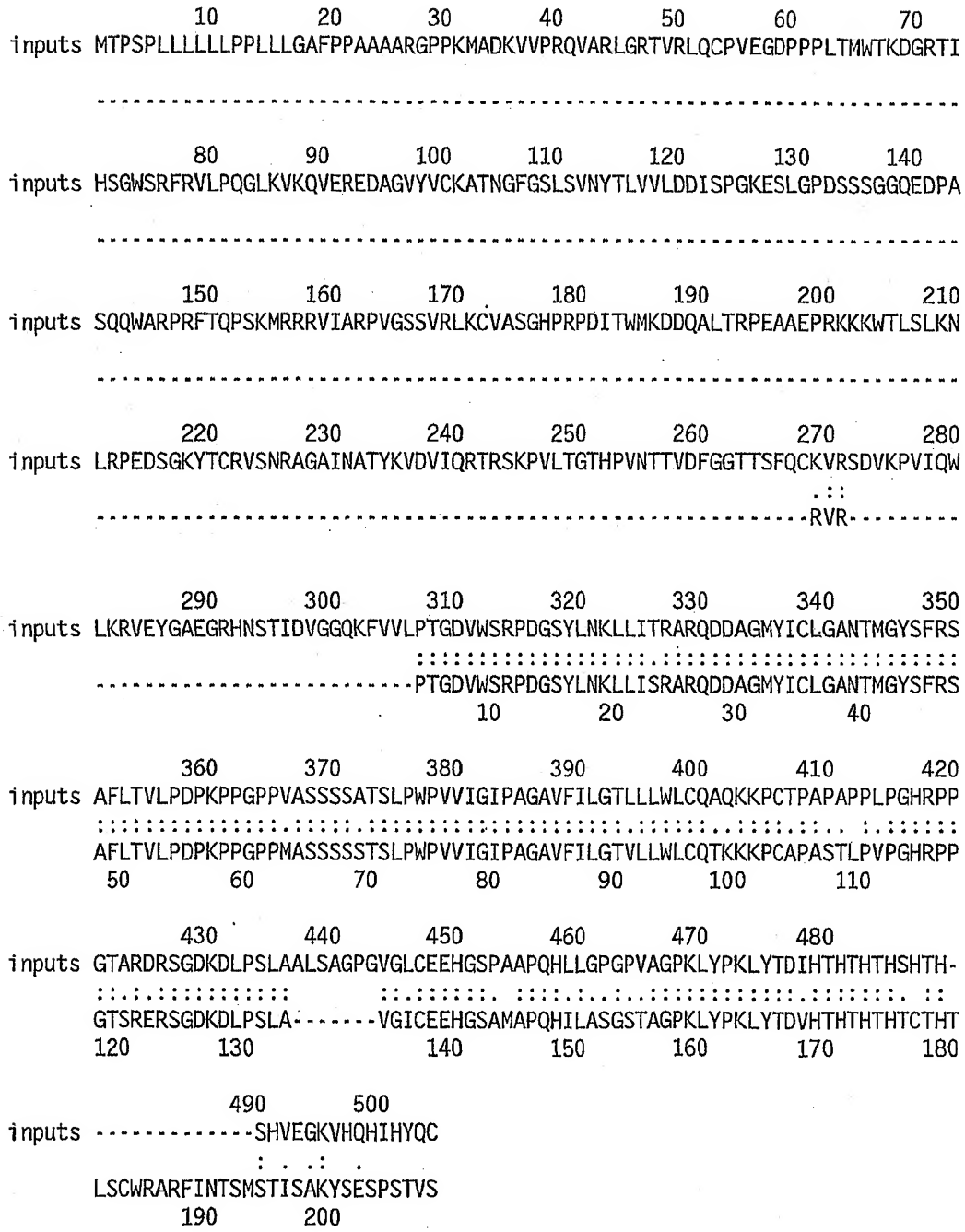

10 20 30 40 50 60 70
inputs MTPSPLLLLLLPPLLLGAFPPAAAARGPPKMADKVVPRQVARLGRTVRLQCPVEGDPPPLTMWTKDGRTI
80 90 100 110 120 130 140
inputs HSGWSRFRVLPQGLKVKQVEREDAGVYVCKATNGFGSLSVNYTLVVLDDISPGKESLGPDSSSGGQEDPA
150 160 170 180 190 200 210
inputs SQQWARPRFTQPSKMRRRVIARPVGSSVRLKCVASGHPRPDITWMKDDQALTRPEAAEPRKKKWTLSLKN
220 230 240 250 260 270 280
inputs LRPEDSGKYTCRVSNRAGAINATYKVDVIQRTRSKPVLTGTHPVNTTVDFGGTTSFQCKVRSDVKPVIQW
RVR
290 300 310 320 330 340 350
inputs LKRVEYGAEGRHNSTIDVGGQKFVVLPTGDVWSRPDGSYLNKLLITRARQDDAGMYICLGANTMGYSFRS
PTGDVWSRPDGSYLNKLLISRARQDDAGMYICLGANTMGYSFRS
10 20 30 40
360 370 380 390 400 410 420
inputs AFLTVLPDPKPPGPPVASSSSATSLPWPVVIGIPAGAVFILGTLLLWLCQAQKKPCTPAPAPPLPGHRPP
AFLTVLPDPKPPGPPMASSSSSTSLPWPVVIGIPAGAVFILGTVLLWLCQTKKKPCAPASTLPVPGHRPP
50 60 70 80 90 100 110
430 440 450 460 470 480
inputs GTARDRSGDKDLPSLAALSAGPGVGLCEEHGSPAAPQHLLGPGPVAGPKLYPKLYTDIHTHTHTHSHTH-
GTSRERSGDKDLPSLA-------VGICEEHGSAMAPQHILASGSTAGPKLYPKLYTDVHTHTHTHTCTHT
120 130 140 150 160 170 180
490 500
inputs ------------SHVEGKVHQHIHYQC
LSCWRARFINTSMSTISAKYSESPSTVS
190 200

FIG.29

```
inputs GT----------------------------------------------------------------
       :
       ATGTCACCGCCTCTGTGTCCCCTCCTTCTCCTGGCTGTGGGCCTGCGGCTGGCTGGAACTCTCAACCCCA
                10        20        30        40        50        60        70

inputs ----------------------------------------------------------------------

       GTGATCCCAATACCTGCAGCTTCTGGGAAAGCTTCACTACCACCACCAAGGAGTCCCACTCCCGCCCCTT
                80        90       100       110       120       130       140

inputs ----------------------------------------------------------------------

       CAGCCTGCTCCCCTCAGAGCCCTGCGAGCGGCCCTGGGAGGGCCCCCATACTTGCCCCAGCCCACAAACT
               150       160       170       180       190       200       210

inputs ----------------------------------------------------------------------

       CAGAGGAAACTCCTGGCTTCTAGGGATTCATTCTGCATGGTCTGTGTCGGGGCTGGAGTGCAGTGGCGAG
               220       230       240       250       260       270       280

inputs ----------------------------------------------------------------------

       ATCGTAGTGCACTGCAACCTCAAACAGGGAATGCGCTTTCTATGCGCCCTCAGCCCAGAGTGTTGAGTGG
               290       300       310       320       330       340       350

inputs ----------------------------------------------------------------------

       TGCCCCTTCCCTGGCCTCCCCTGGCCACACTGTGGTGGTGAAGACGGACCACCGCCAGCGCCTGCAGTGC
               360       370       380       390       400       410       420

inputs ----------------------------------------------------------------------

       TGCCATGGCTTCTATGAGAGCAGGGGGTTCTGTGTCCCGCTCTGTGCCCAGGAGTGTGTCCATGGCCGTT
               430       440       450       460       470       480       490

inputs ----------------------------------------------------------------------

       GTGTGGCACCCAATCAGTGCCAATGTGTGCCAGGCTGGCGGGGCGACGACTGTTCCAGTGCCCCGAACTG
               500       510       520       530       540       550       560

inputs ----------------------------------------------------------------------

       CCTTCAGCCCTGTACCCCTGGCTACTATGGCCCTGCCTGCCAGTTCCGCTGCCAGTGCCATGGGGCACCC
               570       580       590       600       610       620       630

inputs ----------------------------------------------------------------------

       TGCGATCCCCAGACTGGAGCCTGCTTCTGCCCCGCAGAGAGAACTGGGCCCAGCTGTGACGTGTCCTGTT
               640       650       660       670       680       690       700
```

FIG.30A

```
inputs ----------------------------------------------------------------------
       CCCAGGGCACTTCTGGCTTCTTCTGCCCCAGCACCCATCCTTGCCAAAATGGAGGTGTCTTCCAAACCCC
              710       720       730       740       750       760       770

inputs ----------------------------------------------------------------------
       ACAGGGCTCCTGCAGCTGCCCCCCTGGCTGGATGGGCACCATCTGCTCCCTGCCCTGCCCAGAGGGCTTT
              780       790       800       810       820       830       840

inputs ----------------------------------------------------------------------
       CACGGACCCAACTGCTCCCAGGAATGTCGCTGCCACAACGGCGGCCTCTGTGACCGATTCACTGGGCAGT
              850       860       870       880       890       900       910

inputs ----------------------------------------------------------------------
       GCCGCTGCGCTCCGGGTTACACTGGGGATCGGTGCCGGGAGGAGTGCCCGGTGGGCCGCTTTGGGCAGGA
              920       930       940       950       960       970       980

inputs ----------------------------------------------------------------------
       CTGTGCTGAGACGTGCGACTGCGCCCCGGACGCCCGTTGCTTCCCGGCCAACGGCGCATGTCTGTGCGAA
              990      1000      1010      1020      1030      1040      1050

inputs ----------------------------------------------------------------------
       CACGGCTTCACTGGGGACCGCTGCACGGATCGCCTCTGCCCCGACGGCTTCTACGGTCTCAGCTGCCAGG
             1060      1070      1080      1090      1100      1110      1120

inputs ------------CGACC-----------------------------------------------------
                   :::::
       CCCCCTGCACCTGCGACCGGGAGCACAGCCTCAGCTGCCACCCGATGAACGGGGAGTGCTCCTGCCTGCC
             1130      1140      1150      1160      1170      1180      1190

                                                  10
inputs ----------------------------------------CACGC---------------------
                                               :::::
       GGGCTGGGCGGGCCTCCACTGCAACGAGAGCTGCCCGCAGGACACGCATGGGCCAGGGTGCCAGGAGCAC
             1200      1210      1220      1230      1240      1250      1260

inputs ----------------------------------------------------------------------
       TGTCTCTGCCTGCACGGTGGCGTCTGCCAGGCTACCAGCGGCCTCTGTCAGTGCGCGCCGGGTTACACGG
             1270      1280      1290      1300      1310      1320      1330

inputs ----------------------------------------------------------------------
       GCCCTCACTGTGCTAGTCTTTGTCCTCCTGACACCTACGGTGTCAACTGTTCTGCACGCTGCTCATGTGA
             1340      1350      1360      1370      1380      1390      1400
```

FIG.30B

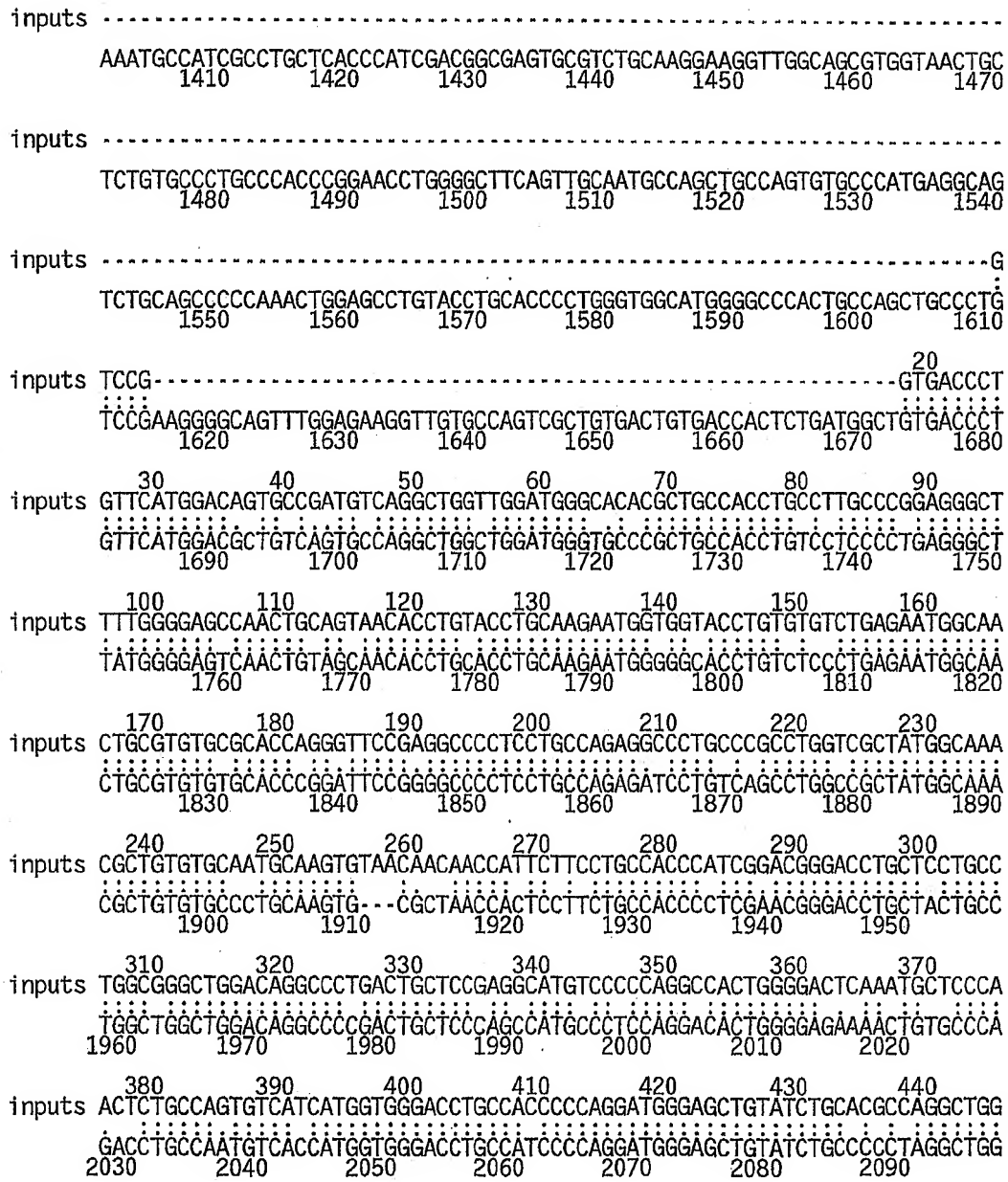


FIG.30C

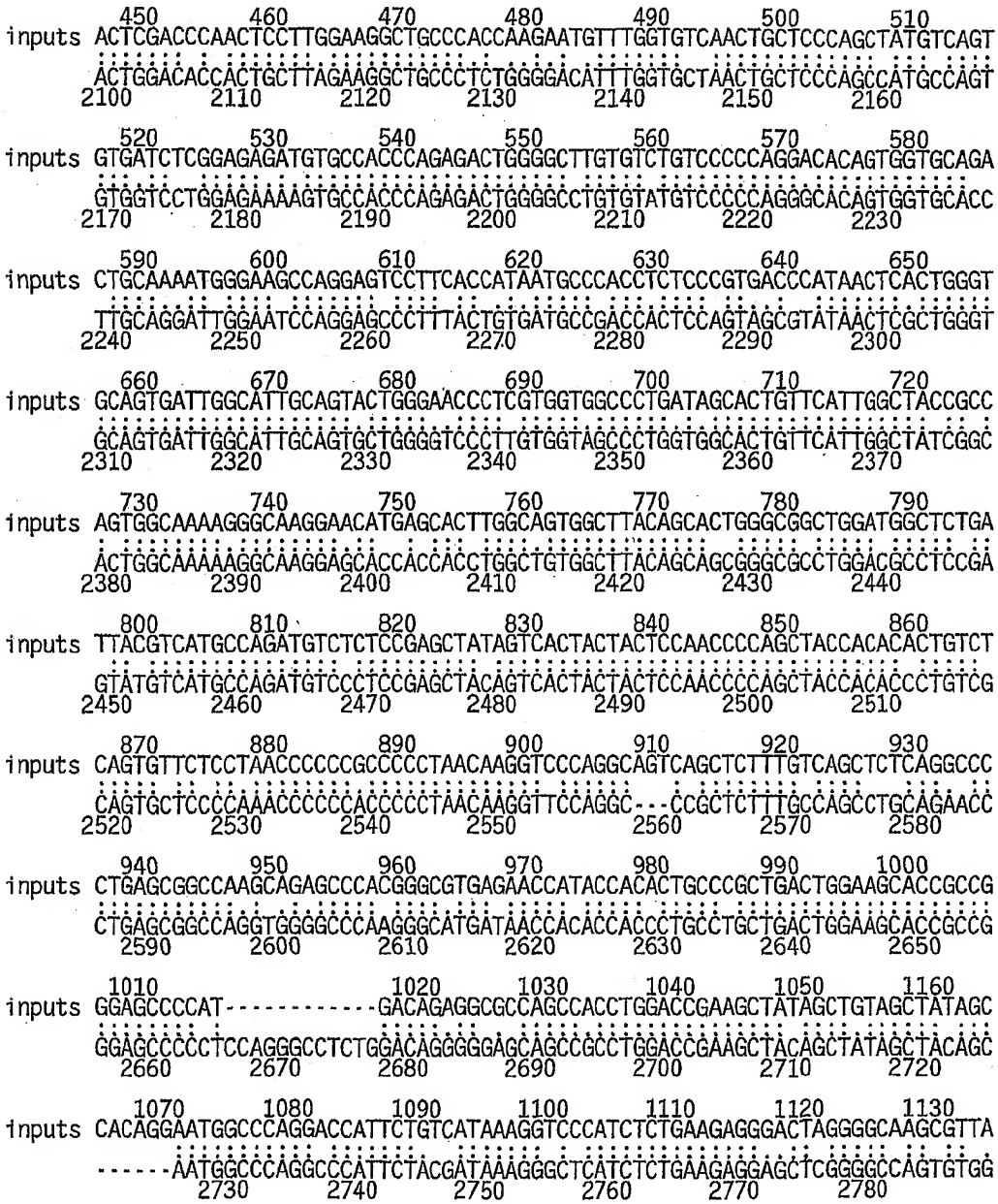

450 460 470 480 490 500 510
inputs ACTCGACCCAACTCCTTGGAAGGCTGCCCACCAAGAATGTTTGGTGTCAACTGCTCCCAGCTATGTCAGT
ACTGGACACCACTGCTTAGAAGGCTGCCCTCTGGGGACATTTGGTGCTAACTGCTCCCAGCCATGCCAGT
2100 2110 2120 2130 2140 2150 2160
520 530 540 550 560 570 580
inputs GTGATCTCGGAGAGATGTGCCACCCAGAGACTGGGGCTTGTGTCTGTCCCCCAGGACACAGTGGTGCAGA
GTGGTCCTGGAGAAAAGTGCCACCCAGAGACTGGGGCCTGTGTATGTCCCCCAGGGCACAGTGGTGCACC
2170 2180 2190 2200 2210 2220 2230
590 600 610 620 630 640 650
inputs CTGCAAAATGGGAAGCCAGGAGTCCTTCACCATAATGCCCACCTCTCCCGTGACCCATAACTCACTGGGT
TTGCAGGATTGGAATCCAGGAGCCCTTTACTGTGATGCCGACCACTCCAGTAGCGTATAACTCGCTGGGT
2240 2250 2260 2270 2280 2290 2300
660 670 680 690 700 710 720
inputs GCAGTGATTGGCATTGCAGTACTGGGAACCCTCGTGGTGGCCCTGATAGCACTGTTCATTGGCTACCGCC
GCAGTGATTGGCATTGCAGTGCTGGGGTCCCTTGTGGTAGCCCTGGTGGCACTGTTCATTGGCTATCGGC
2310 2320 2330 2340 2350 2360 2370
730 740 750 760 770 780 790
inputs AGTGGCAAAAGGGCAAGGAACATGAGCACTTGGCAGTGGCTTACAGCACTGGGCGGCTGGATGGCTCTGA
ACTGGCAAAAAGGCAAGGAGCACCACCACCTGGCTGTGGCTTACAGCAGCGGGCGCCTGGACGCCTCCGA
2380 2390 2400 2410 2420 2430 2440
800 810 820 830 840 850 860
inputs TTACGTCATGCCAGATGTCTCTCCGAGCTATAGTCACTACTACTCCAACCCCAGCTACCACACACTGTCT
GTATGTCATGCCAGATGTCCCTCCGAGCTACAGTCACTACTACTCCAACCCCAGCTACCACACCCTGTCG
2450 2460 2470 2480 2490 2500 2510
870 880 890 900 910 920 930
inputs CAGTGTTCTCCTAACCCCCCGCCCCCTAACAAGGTCCCAGGCAGTCAGCTCTTTGTCAGCTCTCAGGCCC
CAGTGCTCCCCAAACCCCCCACCCCCTAACAAGGTTCCAGGC---CCGCTCTTTGCCAGCCTGCAGAACC
2520 2530 2540 2550 2560 2570 2580
940 950 960 970 980 990 1000
inputs CTGAGCGGCCAAGCAGAGCCCACGGGCGTGAGAACCATACCACACTGCCCGCTGACTGGAAGCACCGCCG
CTGAGCGGCCAGGTGGGGCCCAAGGGCATGATAACCACACCACCCTGCCTGCTGACTGGAAGCACCGCCG
2590 2600 2610 2620 2630 2640 2650
1010 1020 1030 1040 1050 1160
inputs GGAGCCCCAT----------GACAGAGGCGCCAGCCACCTGGACCGAAGCTATAGCTGTAGCTATAGC
GGAGCCCCCTCCAGGGCCTCTGGACAGGGGGAGCAGCCGCCTGGACCGAAGCTACAGCTATAGCTACAGC
2660 2670 2680 2690 2700 2710 2720
1070 1080 1090 1100 1110 1120 1130
inputs CACAGGAATGGCCCAGGACCATTCTGTCATAAAGGTCCCATCTCTGAAGAGGGACTAGGGGCAAGCGTTA
------AATGGCCCAGGCCCATTCTACGATAAAGGGCTCATCTCTGAAGAGGAGCTCGGGGCCAGTGTGG
2730 2740 2750 2760 2770 2780

FIG.30D

```
          1140       1150       1160       1170       1180       1190       1200
inputs TGTCCCTGAGCAGTGAGAACCCCTATGCTACCATCCGAGACCTGCCCAGCCTGCCTGGGGAACCCCGAGA
       :::::::::::::::::::: ::::: :::::::.:::::::::::: ::::.::::. ::::::.::
       CTTCCCTGAGCAGTGAGAACCCATATGCCACCATCCGGGACCTGCCCAGCTTGCCAGGGGGCCCCCGGGA
     2790       2800       2810       2820       2830       2840       2850

          1210       1220       1230       1240       1250       1260       1270
inputs AAGTGGCTATGTGGAGATGAAAGGACCTCCATCAGTGTCCCCTCCCAGGCAGTCTCTTCATCTCCGGGAC
       .:: .:::: .:::::::::::::: ::::: :::: .::  : ::::::::: ::: :::  :  :::::
       GAGCAGCTACATGGAGATGAAAGGCCCTCCCTCAGGATCTGCCCCCAGGCAGCCTCCTCAGTTTTGGGAC
     2860       2870       2880       2890       2900       2910       2920

               1280       1290       1300       1310       1320       1330       1340
inputs AGGCAG---CAGCGGCAACTGCAGCCACAGAGGGACAGCGGCACCTATGAGCAGCCCAGCCCCTTGAGCC
       :: :::   :.:::::::: ::::::::::.::::: :::::::: :::::::::::::: ::: ::
       AGCCAGAGGCGGCGGCAACCCCAGCCACAGAGAGACAGTGGCACCTACGAGCAGCCCAGCCCCCTGATCC
     2930       2940       2950       2960       2970       2980       2990

               1350       1360       1370       1380       1390       1400       1410
inputs ATAATGAAGAGTCTTTGGGCTCCACGCCCCCGCTTCCTCCAGGCCTGCCTCCTGGTCACTACGACTCCCC
       ::.:  .::: ::: :::::::: :::::: :: :::::.::::::.:: :: :: ::::: ::::: ::
       ATGACCGAGACTCTGTGGGCTCCCAGCCCCCTCTGCCTCCGGGCCTACCCCCCGGCCACTATGACTCACC
     3000       3010       3020       3030       3040       3050       3060

               1420       1430       1440       1450       1460       1470       1480
inputs CAAGAACAGCCATATCCCTGGACACTATGACTTGCCTCCAGTACGGCATCCTCCATCCCCTCCATCCCGG
       :::::::::::: :::::::::::: :::::::::::::::::::::::::::: ::::: ::::::    ::.
       CAAGAACAGCCACATCCCTGGACATTATGACTTGCCTCCAGTACGGCATCCCCCATCAGCTCCACTTCGA
     3070       3080       3090       3100       3110       3120       3130

               1490
inputs CGCCAGGACCGC
       ::::::::::::
       CGCCAGGACCGT
     3140       3150
```

FIG.30E

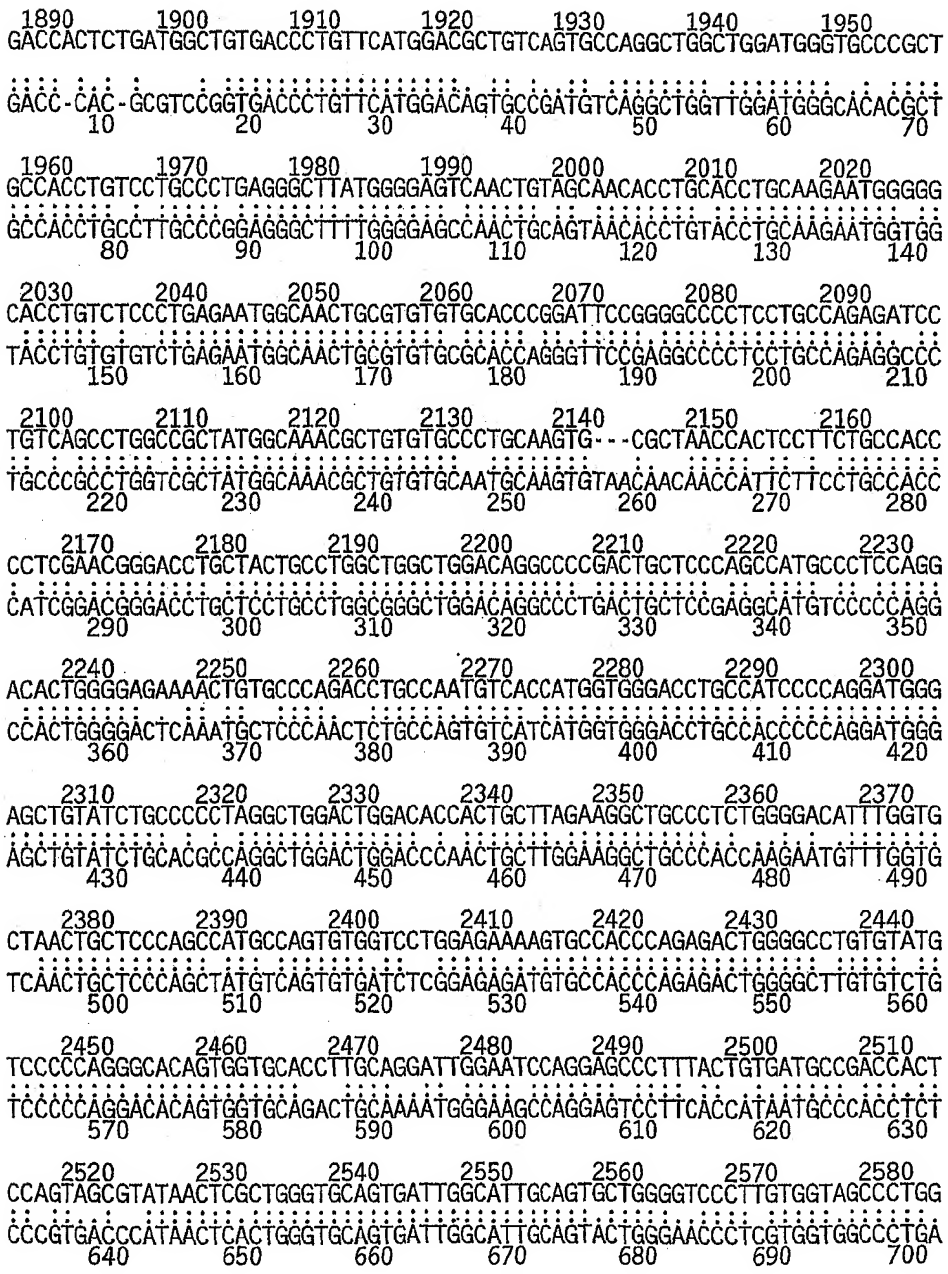


FIG.31A

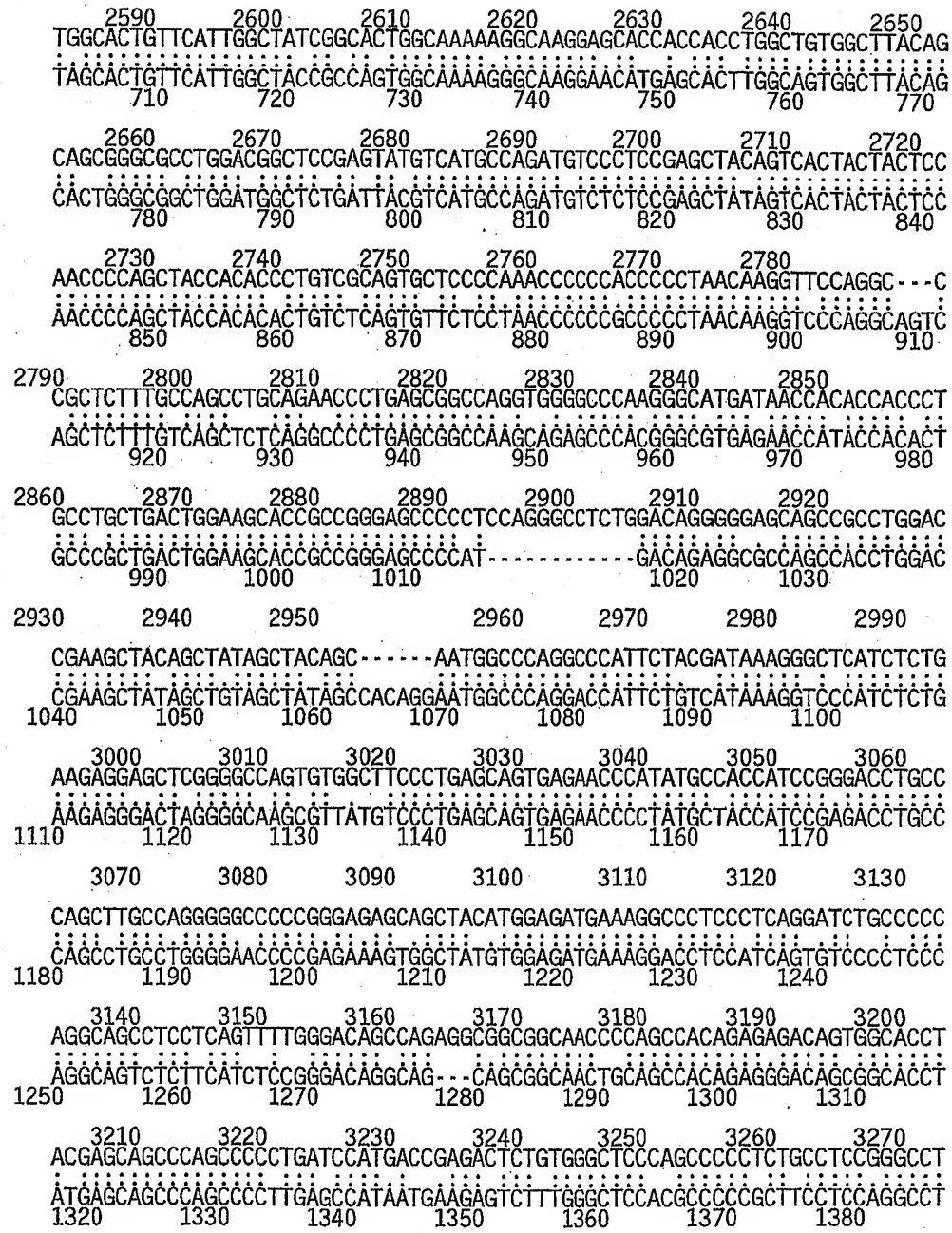


FIG.31B

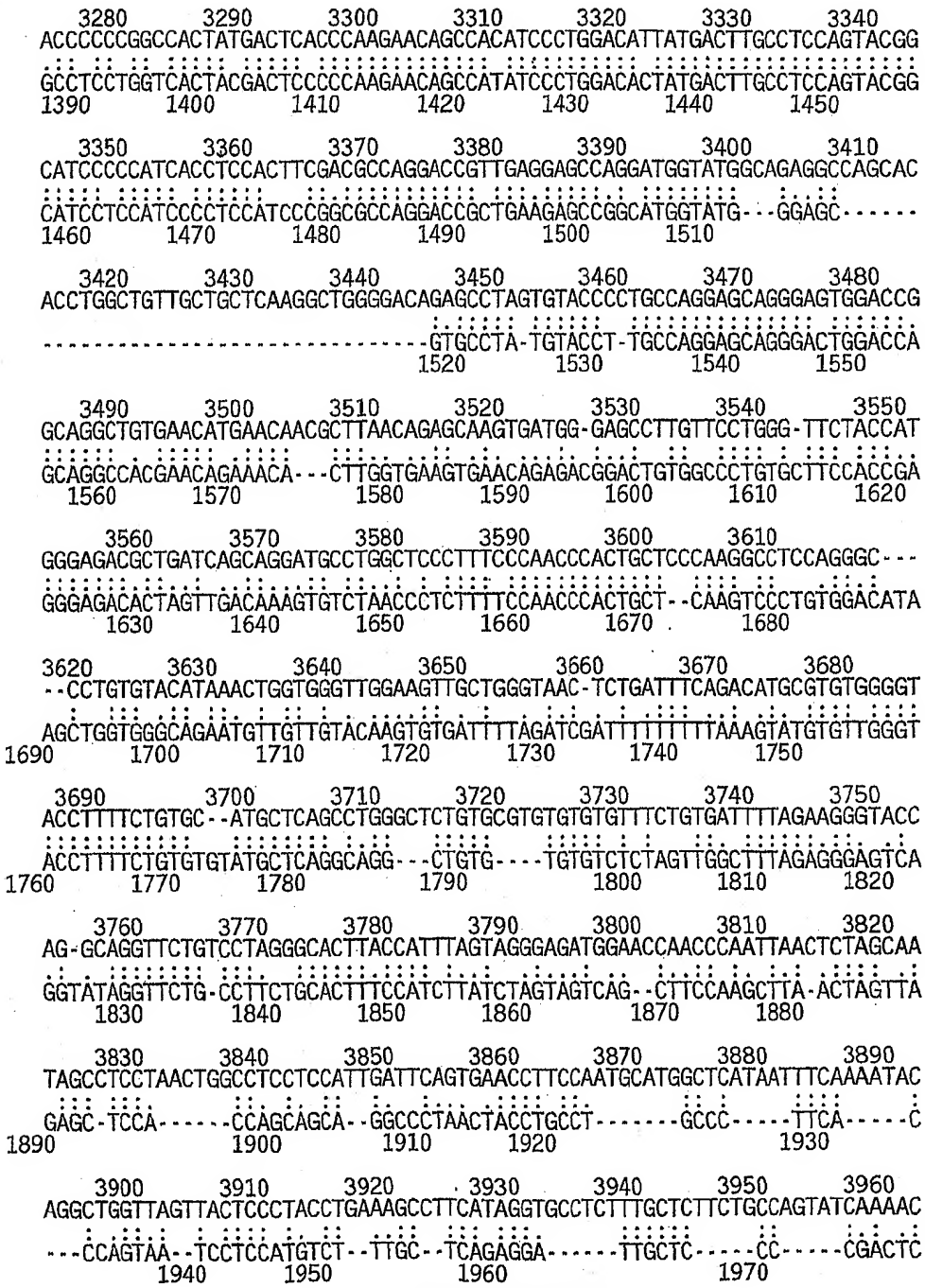


FIG 31C

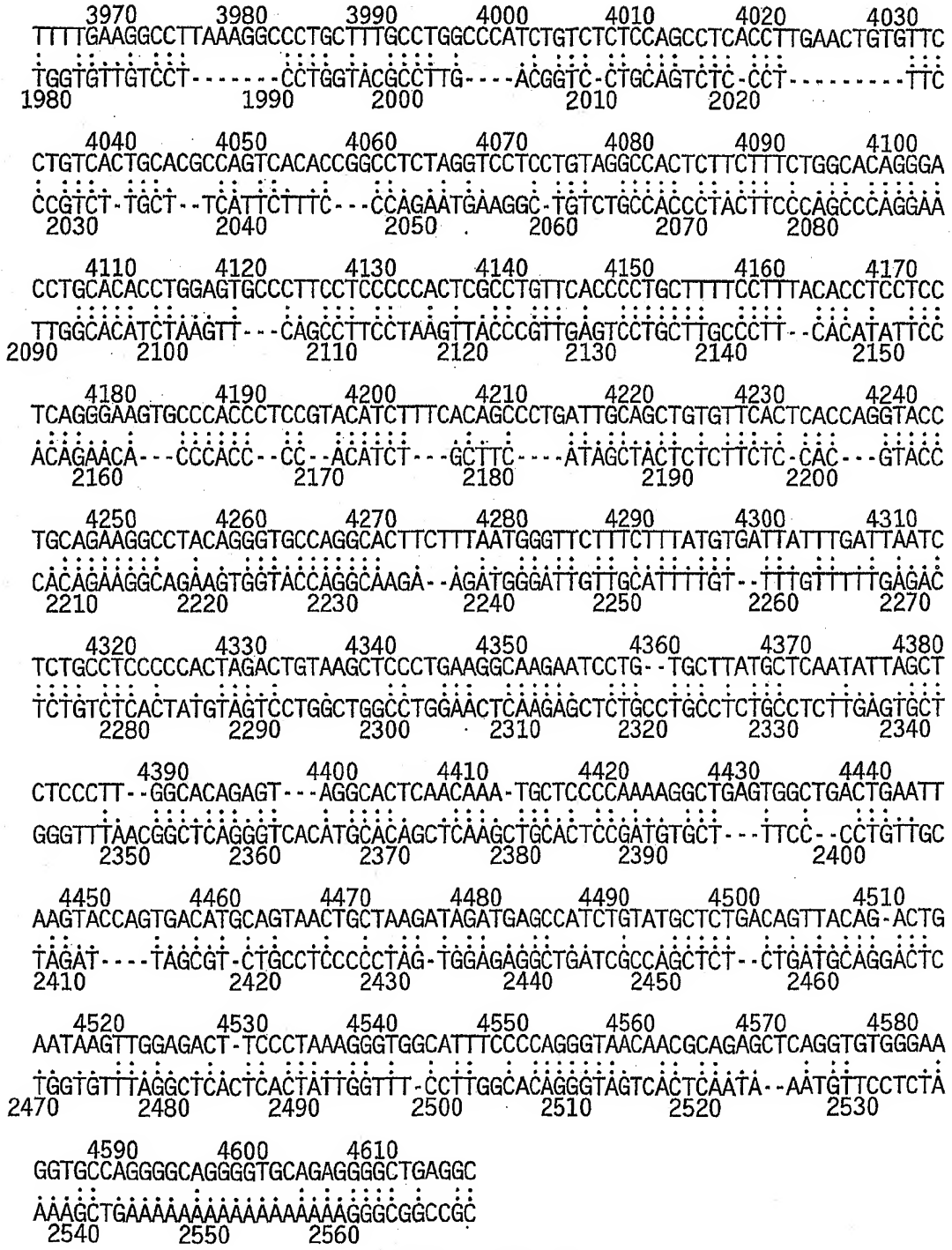

3970 3980 3990 4000 4010 4020 4030
TTTTGAAGGCCTTAAAGGCCCTGCTTTGCCTGGCCCATCTGTCTCTCCAGCCTCACCTTGAACTGTGTTC
TGGTGTTGTCCT-------CCTGGTACGCCTTG----ACGGTC-CTGCAGTCTC-CCT---------TTC
1980 1990 2000 2010 2020
4040 4050 4060 4070 4080 4090 4100
CTGTCACTGCACGCCAGTCACACCGGCCTCTAGGTCCTCCTGTAGGCCACTCTTCTTTCTGGCACAGGGA
CCGTCT-TGCT--TCATTCTTTC---CCAGAATGAAGGC-TGTCTGCCACCCTACTTCCCAGCCCAGGAA
2030 2040 2050 2060 2070 2080
4110 4120 4130 4140 4150 4160 4170
CCTGCACACCTGGAGTGCCCTTCCTCCCCCACTCGCCTGTTCACCCCTGCTTTTCCTTTACACCTCCTCC
TTGGCACATCTAAGTT---CAGCCTTCCTAAGTTACCCGTTGAGTCCTGCTTGCCCTT--CACATATTCC
2090 2100 2110 2120 2130 2140 2150
4180 4190 4200 4210 4220 4230 4240
TCAGGGAAGTGCCCACCCTCCGTACATCTTTCACAGCCCTGATTGCAGCTGTGTTCACTCACCAGGTACC
ACAGAACA---CCCACC--CC--ACATCT---GCTTC----ATAGCTACTCTCTTCTC-CAC---GTACC
2160 2170 2180 2190 2200
4250 4260 4270 4280 4290 4300 4310
TGCAGAAGGCCTACAGGGTGCCAGGCACTTCTTTAATGGGTTCTTTCTTTATGTGATTATTTGATTAATC
CACAGAAGGCAGAAGTGGTACCAGGCAAGA--AGATGGGATTGTTGCATTTTGT--TTTGTTTTTGAGAC
2210 2220 2230 2240 2250 2260 2270
4320 4330 4340 4350 4360 4370 4380
TCTGCCTCCCCCACTAGACTGTAAGCTCCCTGAAGGCAAGAATCCTG--TGCTTATGCTCAATATTAGCT
TCTGTCTCACTATGTAGTCCTGGCTGGCCTGGAACTCAAGAGCTCTGCCTGCCTCTGCCTCTTGAGTGCT
2280 2290 2300 2310 2320 2330 2340
4390 4400 4410 4420 4430 4440
CTCCCTT--GGCACAGAGT---AGGCACTCAACAAA-TGCTCCCCAAAAGGCTGAGTGGCTGACTGAATT
GGGTTTAACGGCTCAGGGTCACATGCACAGCTCAAGCTGCACTCCGATGTGCT---TTCC--CCTGTTGC
2350 2360 2370 2380 2390 2400
4450 4460 4470 4480 4490 4500 4510
AAGTACCAGTGACATGCAGTAACTGCTAAGATAGATGAGCCATCTGTATGCTCTGACAGTTACAG-ACTG
TAGAT----TAGCGT-CTGCCTCCCCCTAG-TGGAGAGGCTGATCGCCAGCTCT--CTGATGCAGGACTC
2410 2420 2430 2440 2450 2460
4520 4530 4540 4550 4560 4570 4580
AATAAGTTGGAGACT-TCCCTAAAGGGTGGCATTTCCCCAGGGTAACAACGCAGAGCTCAGGTGTGGGAA
TGGTGTTTAGGCTCACTCACTATTGGTTT-CCTTGGCACAGGGTAGTCACTCAATA--AATGTTCCTCTA
2470 2480 2490 2500 2510 2520 2530
4590 4600 4610
GGTGCCAGGGGCAGGGGTGCAGAGGGGCTGAGGC
AAAGCTGAAAAAAAAAAAAAAAAAGGGCGGCCGC
2540 2550 2560

FIG. 31D

```
                10        20        30        40        50        60        70
inputs MSPPLCPLLLLAVGLRLAGTLNPSDPNTCSFWESFTTTTKESHSRPFSLLPSEPCERPWEGPHTCPSPQT

       ----------------------------------------------------------------------

                80        90       100       110       120       130       140
inputs QRKLLASRDSFCMVCVGAGVQWRDRSALQPQTGNALSMRPQPRVLSGAPSLASPGHTVVVKTDHRQRLQC

       ----------------------------------------------------------------------

               150       160       170       180       190       200       210
inputs CHGFYESRGFCVPLCAQECVHGRCVAPNQCQCVPGWRGDDCSSAPNCLQPCTPGYYGPACQFRCQCHGAP

       ----------------------------------------------------------------------

               220       230       240       250       260       270       280
inputs CDPQTGACFCPAERTGPSCDVSCSQGTSGFFCPSTHPCQNGGVFQTPQGSCSCPPGWMGTICSLPCPEGF

       ----------------------------------------------------------------------

               290       300       310       320       330       340       350
inputs HGPNCSQECRCHNGGLCDRFTGQCRCAPGYTGDRCREECPVGRFGQDCAETCDCAPDARCFPANGACLCE

       ----------------------------------------------------------------------

               360       370       380       390       400       410       420
inputs HGFTGDRCTDRLCPDGFYGLSCQAPCTCDREHSLSCHPMNGECSCLPGWAGLHCNESCPQDTHGPGCQEH
                                                                   .::..:
       ------------------------------------------------------------STHASG----

               430       440       450       460       470       480       490
inputs CLCLHGGVCQATSGLCQCAPGYTGPHCASLCPPDTYGVNCSARCSCENAIACSPIDGECVCKEGWQRGNC

       ----------------------------------------------------------------------

               500       510       520       530       540       550       560
inputs SVPCPPGTWGFSCNASCQCAHEAVCSPQTGACTCTPGWHGAHCQLPCPKGQFGEGCASRCDCDHSDGCDP
                                                                            ::
       --------------------------------------------------------------------DP

               570       580       590       600       610       620       630
inputs VHGRCQCQAGWMGARCHLSCPEGLWGVNCSNTCTCKNGGTCLPENGNCVCAPGFRGPSCQRSCQPGRYGK
       :::.:.:::::::.::::.::::.::.::::::::::::::..::::::::::::::::::.:.::::::
       VHGQCRCQAGWMGTRCHLPCPEGFWGANCSNTCTCKNGGTCVSENGNCVCAPGFRGPSCQRPCPPGRYGK
       10        20        30        40        50        60        70
```

FIG.32A

```
                640        650        660        670        680        690
inputs RCVPCKCAN-HSFCHPSNGTCYCLAGWTGPDCSQPCPPGHWGENCAQTCQCHHGGTCHPQDGSCICPLGW
       :::.:::.: :: ::::.::: ::::::::::::..:::::::: .:.: ::::::::::::::::::  ::
       RCVQCKCNNNHSSCHPSDGTCSCLAGWTGPDCSEACPPGHWGLKCSQLCQCHHGGTCHPQDGSCICTPGW
       80         90         100        110        120        130        140

    700        710        720        730        740        750        760
inputs TGHHCLEGCPLGTFGANCSQPCQCGPGEKCHPETGACVCPPGHSGAPCRIGIQEPFTVMPTTPVAYNSLG
       :: .:::::::   ::.:::: :::. ::.:::::::::::::::::: :..: ::.::.:::.::. :::: 
       TGPNCLEGCPPRMFGVNCSQLCQCDLGEMCHPETGACVCPPGHSGADCKMGSQESFTIMPTSPVTHNSLG
      150        160        170        180        190        200        210

    770        780        790        800        810        820        830
inputs AVIGIAVLGSLVVALVALFIGYRHWQKGKEHHHLAVAYSSGRLDGSEYVMPDVPPSYSHYYSNPSYHTLS
       ::::::::::.:::::.::::::::.:::::::: :::::::.:::::::.:::::.:::::::::::::::::
       AVIGIAVLGTLVVALIALFIGYRQWQKGKEHEHLAVAYSTGRLDGSDYVMPDVSPSYSHYYSNPSYHTLS
      220        230        240        250        260        270        280

    840        850        860        870        880        890        900
inputs QCSPNPPPPNKVPGP-LFASLQNPERPGGAQGHDNHTTLPADWKHRREPPPGPLDRGSSRLDRSYSYSYS
       ::::::::::::::. ::.: :.:::::. :.:..:::::::::::::::::     :::.:.:::::: :::.
       QCSPNPPPPNKVPGSQLFVSSQAPERPSRAHGRENHTTLPADWKHRREPH----DRGASHLDRSYSCSYS
      290        300        310        320        330            340        350

       910        920        930        940        950        960        970
inputs --NGPGPFYDKGLISEEELGASVASLSSENPYATIRDLPSLPGGPRESSYMEMKGPPSGSAPRQPPQFWD
         :::::: .:: :::: ::::: :::::::::::::::::: ::::.:.::::::: :.:::. ...:
       HRNGPGPFCHKGPISEEGLGASVMSLSSENPYATIRDLPSLPGEPRESGYVEMKGPPSVSPPRQSLHLRD
         360        370        380        390        400        410        420

       980        990       1000       1010       1020       1030       1040
inputs SQRRRQPQPQRDSGTYEQPSPLIHDRDSVGSQPPLPPGLPPGHYDSPKNSHIPGHYDLPPVRHPPSPPLR
       :.: : ::::::::::::::: :. .:.:: :::::::::::::::::::::::::::::::::::: :
       RQQR-QLQPQRDSGTYEQPSPLSHNEESLGSTPPLPPGLPPGHYDSPKNSHIPGHYDLPPVRHPPSPPSR
          430        440        450        460        470        480        490

       1050
inputs RQDR
       ::::
       RQDR
```

FIG.32B

Input file T272Atrxa6b6; Output File T272Atrxa6b6.pat
Sequence length 3567

```
GTCCGACCCACGCGTCCGAGCCACACCCTGAAGGTGGTTGGAAGGAGGGAAGGATCTAGGTCCTGAGCACTGGAATTCC  79

CCAGAACAGCATCTGGCTTCCCAGACCCATGCTGGCCACCACTGATGTGTCCTTCCGGCTGCTGGCTGCAGTGCTGTTC 158

TGTTGTTGGGTGCCCTGTGGCAGGCTTGTGCAATGCCACTCTGTCCCCTCCTCCTCCTGGCCCTAGGCCTGCGTCTGGC 237

TGGAACACTCAACTCCAATGATCCCAATGTCTGTACCTTCTGGGAAAGCTTCACCACGACCACTAAGGAGTCCCACCTT 316

CGCCCCTTCAGCCTGCCCCCAGCCGAGTCCTGCGACAGGCCCTGGGAAGACCCCCACACCTGCGCTCAGCCTACGGTTG 395

TCTACCGGACTGTGTACCGTCAGGTGGTGAAGATGGACTCCCGCCCACGCCTGCAGTGCTGTGGGGGTTACTACGAGAG 474

CAGTGGAGCCTGTGTCCCACTCTGTGCCCAGGAGTGTGTCCACGGTCGCTGTGTGGCTCCTAATCGGTGCCAGTGTGCA 553

CCAGGCTGGCGGGGTGACGACTGTTCCAGTGAGTGTGCTCCTGGAATGTGGGGACCACAGTGTGACAGGCTCTGCCTCT 632

GTGGCAACAGCAGTTCCTGTGATCCCAGGAGTGGGGTGTGTTTTTGCCCCTCTGGCCTGCAGCCCCCCGACTGCCTTCA 711

GCCTTGCCCCGATGGCCACTATGGTCCTGCCTGCCAGTTTGATTGCCATTGCTATGGGGCATCCTGTGACCCCCGGGAT 790

GGAGCCTGCTTCTGCCCCCCAGGGAGAACAGGACCCAGGGCACTGATGGCTTCTTCTGCCCCAGAACTTATCCTTGCCA 869

                                                    M  G  V  I  C  S     6
AAATGGAGGTGTTCCTCAGGGCTCTCAAGGCTCCTGCAGCTGCCCACCGGGCTGG ATG GGT GTC ATC TGT TCC  942

 L  P  C  P  E  G  F  H  G  P  N  C  T  Q  E  C  R  C  H  N    26
CTG CCA TGC CCA GAG GGT TTC CAC GGA CCC AAC TGT ACT CAG GAA TGT CGT TGC CAC AAT 1002

 G  G  L  C  D  R  F  T  G  Q  C  H  C  A  P  G  Y  I  G  D    46
GGT GGC CTT TGT GAC AGG TTT ACT GGG CAG TGC CAC TGT GCT CCT GGC TAT ATC GGG GAT 1062

 R  C  R  E  E  C  P  V  G  R  F  G  Q  D  C  A  E  T  C  D    66
CGG TGC CGT GAA GAG TGC CCT GTG GGC CGC TTC GGT CAA GAC TGT GCT GAG ACC TGT GAC 1122

 C  A  P  G  A  R  C  F  P  A  N  G  A  C  L  C  E  H  G  F    86
TGT GCT CCT GGC GCT CGT TGC TTT CCT GCC AAT GGC GCG TGT CTG TGC GAA CAT GGC TTC 1182
```

FIG.33A

```
 T  G  D  R  C  T  E  R  L  C  P  D  G  R  Y  G  L  S  C  Q   106
ACA GGC GAC CGC TGC ACT GAG CGA CTC TGT CCA GAT GGC CGC TAT GGT CTG AGC TGC CAA  1242

 D  P  C  T  C  D  P  E  H  S  L  S  C  H  P  M  H  G  E  C   126
GAT CCC TGC ACC TGC GAC CCA GAA CAC AGT CTC AGC TGC CAC CCA ATG CAC GGC GAG TGC  1302

 S  C  Q  P  G  W  A  G  L  H  C  N  E  S  C  P  Q  D  T  H   146
TCC TGC CAG CCA GGT TGG GCG GGC CTC CAC TGC AAC GAG AGC TGC CCT CAG GAC ACG CAC  1362

 G  A  G  C  Q  E  H  C  L  C  L  H  G  G  V  C  L  A  D  S   166
GGA GCC GGT TGC CAG GAG CAC TGC CTC TGT CTG CAC GGC GGT GTT TGC CTC GCC GAC AGC  1422

 G  L  C  R  C  A  P  G  Y  T  G  P  H  C  A  N  L  C  P  P   186
GGC CTC TGC CGG TGT GCA CCT GGC TAC ACG GGA CCT CAC TGC GCT AAT CTT TGT CCA CCT  1482

 N  T  Y  G  I  N  C  S  S  H  C  S  C  E  N  A  I  A  C  S   206
AAC ACT TAT GGG ATC AAC TGT TCC TCC CAC TGC TCC TGT GAA AAT GCC ATT GCC TGC TCT  1542

 P  V  D  G  T  C  I  C  K  E  G  W  Q  R  G  N  C  S  V  P   226
CCT GTC GAC GGC ACG TGC ATC TGC AAG GAA GGT TGG CAG CGT GGT AAC TGC TCT GTG CCC  1602

 C  P  P  G  T  W  G  F  S  C  N  A  S  C  Q  C  A  H  E  G   246
TGT CCC CCT GGC ACC TGG GGC TTC AGT TGC AAT GCC AGT TGC CAG TGT GCC CAC GAG GGA  1662

 V  C  S  P  Q  T  G  A  C  T  C  T  P  G  W  R  G  V  H  C   266
GTC TGC AGC CCC CAA ACT GGA GCC TGT ACT TGC ACC CCT GGG TGG CGT GGG GTT CAC TGC  1722

 Q  L  P  C  P  K  G  Q  F  G  E  G  C  A  S  V  C  D  C  D   286
CAA CTT CCG TGC CCG AAG GGA CAG TTT GGT GAA GGT TGT GCC AGT GTC TGT GAC TGT GAC  1782

 H  S  D  G  C  D  P  V  H  G  H  C  R  C  Q  A  G  W  M  G   306
CAC TCC GAT GGC TGT GAC CCT GTT CAT GGA CAC TGC CGA TGT CAG GCT GGC TGG ATG GGC  1842

 T  R  C  H  L  P  C  P  E  G  F  W  G  A  N  C  S  N  A  C   326
ACA CGT TGC CAC CTG CCT TGC CCA GAG GGC TTT TGG GGA GCC AAC TGC AGC AAT GCC TGT  1902

 T  C  K  N  G  G  T  C  V  P  E  N  G  N  C  V  C  A  P  G   346
ACC TGC AAG AAT GGT GGC ACT TGT GTA CCT GAG AAC GGC AAC TGT GTG TGC GCA CCA GGG  1962
```

FIG.33B

```
 F  R  G  P  S  C  Q  R  P  C  P  P  G  R  Y  G  K  R  C  V   366
TTC AGA GGC CCC TCC TGC CAG AGG CCC TGC CCG CCT GGT CGC TAT GGC AAA CGC TGT GTG  2022

 P  C  K  C  N  N  H  S  S  C  H  P  S  D  G  T  C  S  C  L   386
CCC TGC AAG TGC AAC AAC CAT TCT TCC TGC CAC CCG TCG GAT GGG ACC TGC TCC TGC CTG  2082

 A  G  W  T  G  P  D  C  S  E  S  C  P  P  G  H  W  G  L  K   406
GCA GGC TGG ACA GGC CCT GAC TGC TCT GAA TCA TGT CCC CCA GGC CAC TGG GGA CTC AAA  2142

 C  S  Q  P  C  Q  C  H  H  G  A  T  C  H  P  Q  D  G  S  C   426
TGC TCC CAA CCC TGC CAG TGT CAT CAT GGT GCC ACC TGC CAC CCC CAG GAT GGG AGC TGT  2202

 V  C  I  P  G  W  T  G  P  N  C  S  E  G  C  P  S  R  M  F   446
GTC TGC ATC CCA GGC TGG ACT GGA CCC AAC TGC TCG GAA GGC TGC CCA TCA AGA ATG TTT  2262

 G  V  N  C  S  Q  L  C  Q  C  D  P  G  E  M  C  H  P  E  T   466
GGT GTC AAC TGC TCC CAG CTA TGT CAG TGT GAT CCT GGA GAG ATG TGC CAC CCA GAG ACT  2322

 G  A  C  V  C  P  P  G  H  S  G  A  H  C  K  V  G  S  Q  E   486
GGG GCT TGC GTC TGT CCC CCA GGA CAC AGT GGT GCG CAC TGC AAA GTG GGC AGC CAG GAG  2382

 S  F  T  I  M  P  T  S  P  V  I  H  N  S  L  G  A  V  I  G   506
TCC TTC ACC ATA ATG CCC ACC TCT CCT GTG ATC CAT AAC TCA CTG GGT GCC GTG ATT GGC  2442

 I  A  V  L  G  T  L  V  V  A  L  V  A  L  F  I  G  Y  R  H   526
ATT GCA GTG CTG GGG ACC CTT GTG GTG GCC CTG GTA GCA CTG TTT ATT GGC TAC CGA CAC  2502

 W  Q  K  G  K  E  H  E  H  L  A  V  A  Y  S  T  G  R  L  D   546
TGG CAA AAG GGC AAG GAA CAT GAG CAC TTG GCA GTG GCT TAC AGC ACT GGG CGA CTG GAT  2562

 G  S  D  Y  V  M  P  D  V  S  P  S  Y  S  H  Y  Y  S  N  P   566
GGC TCC GAT TAC GTC ATG CCA GAT GTC TCT CCG AGC TAC AGT CAC TAC TAT TCC AAC CCT  2622

 S  Y  H  T  L  S  Q  C  S  P  N  P  P  P  P  N  K  I  P  G   586
AGC TAC CAC ACA CTG TCT CAG TGT TCT CCT AAC CCT CCA CCC CCT AAC AAG ATT CCA GGC  2682

 S  Q  L  F  V  S  S  Q  A  S  E  R  P  N  R  N  H  G  R  D   606
AGT CAG CTG TTT GTC AGC TCC CAG GCA TCT GAG CGG CCA AAC AGA AAC CAT GGG CGA GAT  2742
```

FIG.33C

```
 N  H  A  T  L  P  A  D  W  K  H  R  R  E  S  H  D  R  A  F   626
AAC CAC GCC ACA CTG CCC GCT GAC TGG AAG CAC CGA CGG GAG TCC CAT GAC AGA GCT TTC  2802

 L  R  H  Q  P  P  G  P  K  V  *                                637
CTC AGG CAC CAG CCA CCT GGA CCG AAG GTA TAG                     2835

CTGTAGCTATGGCCACAGGAATGGCCCGGGGCCATTCTGTCATAAAGGTCCCATCTCTGAAGAAGGACTAGGGGCAAGC  2914

GTTATGTCCCTGAGCAGTGAGAACCCCTATGCGACCATCCGAGACCTGCCCGGCCTGCCTGGGGAACCCCGAGAAAGCA  2993

GCTATGTGGAGATGAAAGGCCCTCCATCAGTGTCTCCCCCCAGGCAGCCTCTTCATCTCCGGGACAGGCAGCAGCAGCA  3072

ACTGCAGTCTCAGAGAGACAGCGGCACCTATGAGCAGCCCACTCCCTTGAGCCGTAATGAAGAGTCTGTGGGCTCCATG  3151

CCCCCTCTTCCTCCGGGCCTGCCACCCGGCCACTATGACTCGCCCAAAAACAGCCACATCCCTGGACACTATGACTTGC  3230

CTCCAGTACGGCATCCTCCATCACCTCCATCCCGGCGCCAGGACCGCTGAGGAGCCAGCATGGTATGGGAGAGTGCCTG  3309

TGAACCCTGCCAGGAGCAGGGCCTGGACCAGCAGGCCATGAATAGACATACTTGGTGAAGTGAACGGAGACTGAGGATG  3388

GCTCTGCTTCCACCGAGGGAGACACTAGTTGGCAAAGTGTCTAACCTCCCTTTTCCAGCCCATTGCTCAAGTCCCCCAG  3467

GCTGTGGACATGAGCTGGTGGGCAGAATGTTGTTGTTGAAGTCTGATTTTAGATTGATTTTTTAAAAAAAAAAAAAAAA  3546

AAAAAAAAAAGGGCGGCCGC                                            3567
```

FIG.33D

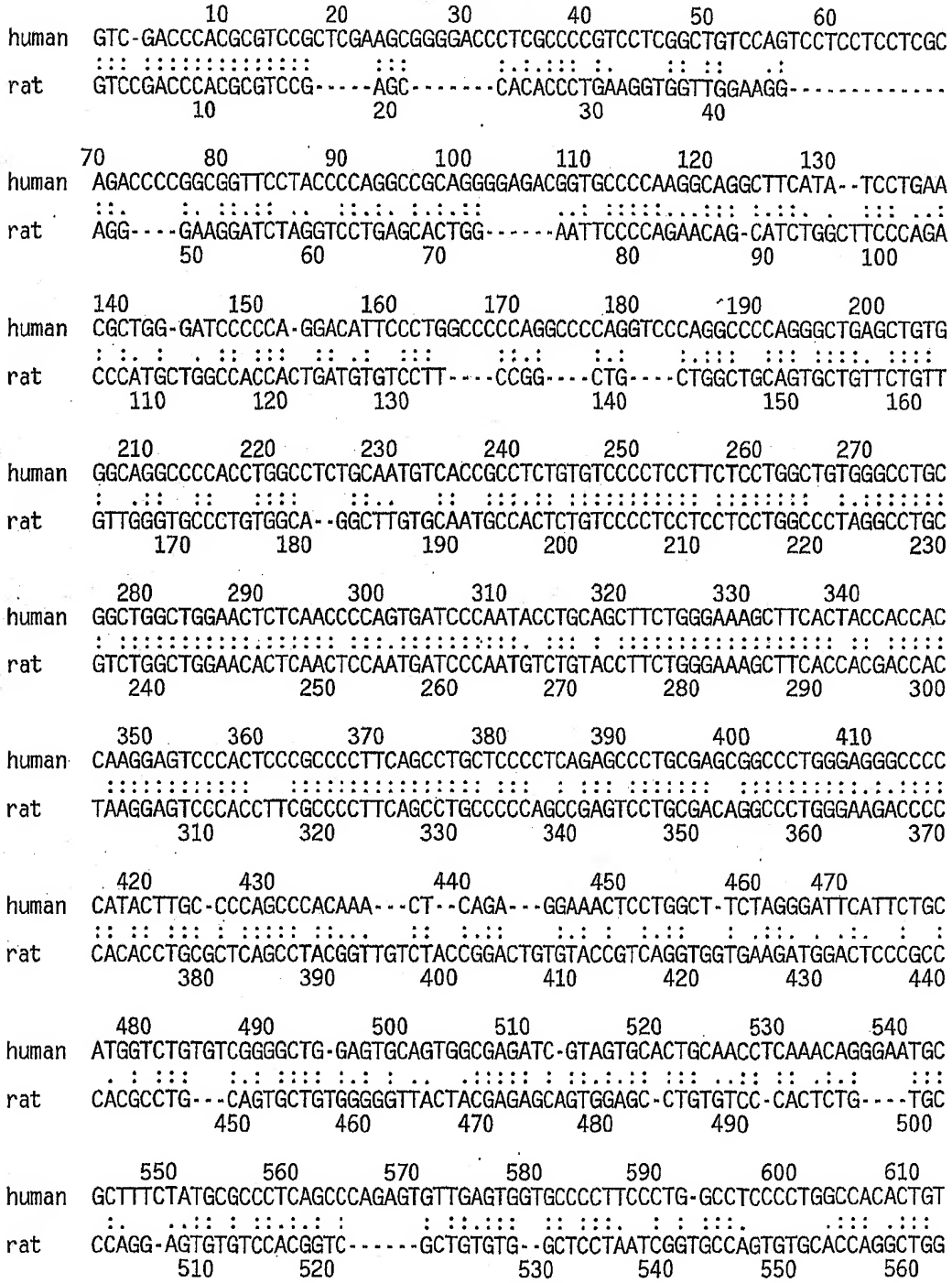


FIG.34A

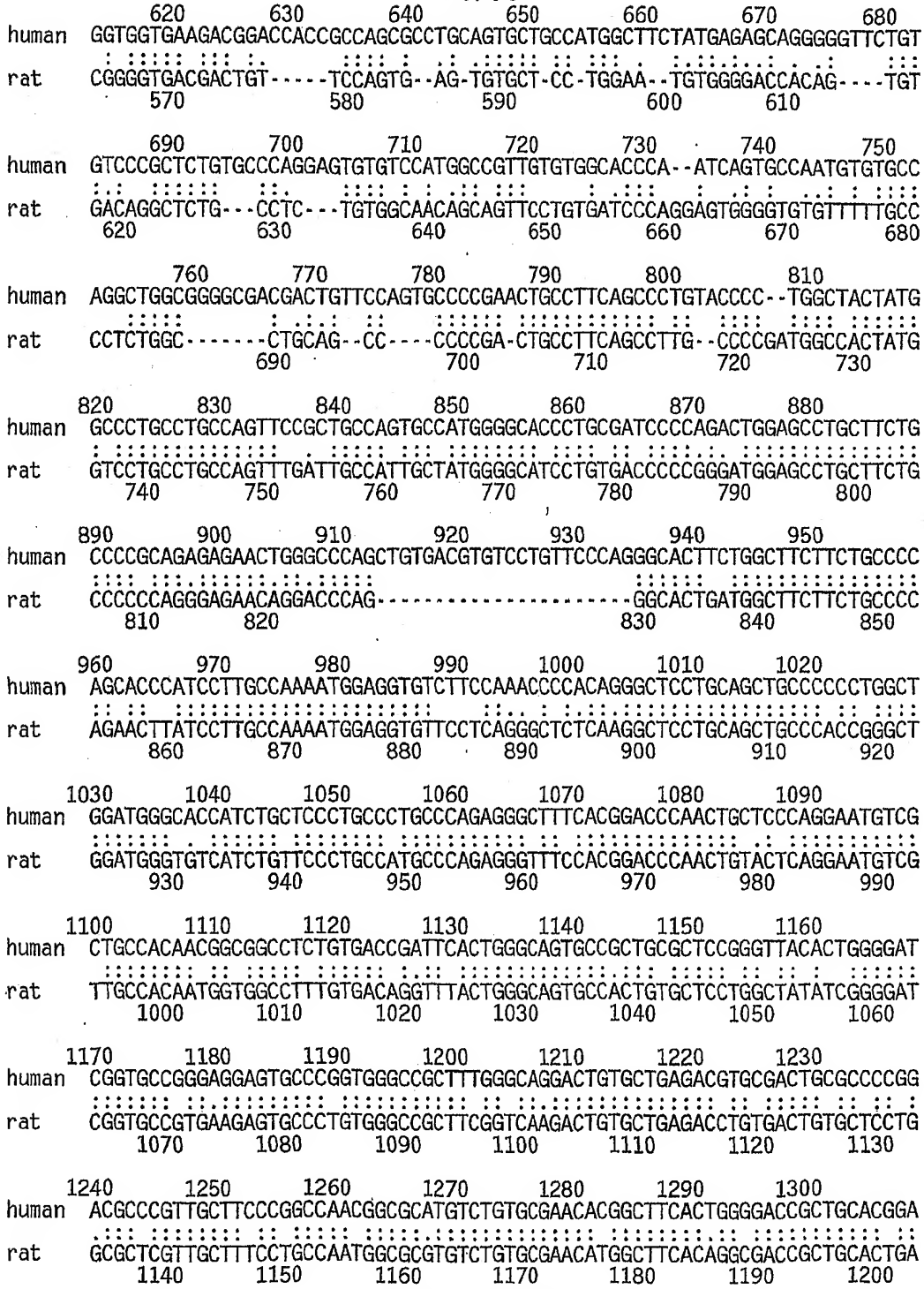


FIG.34B

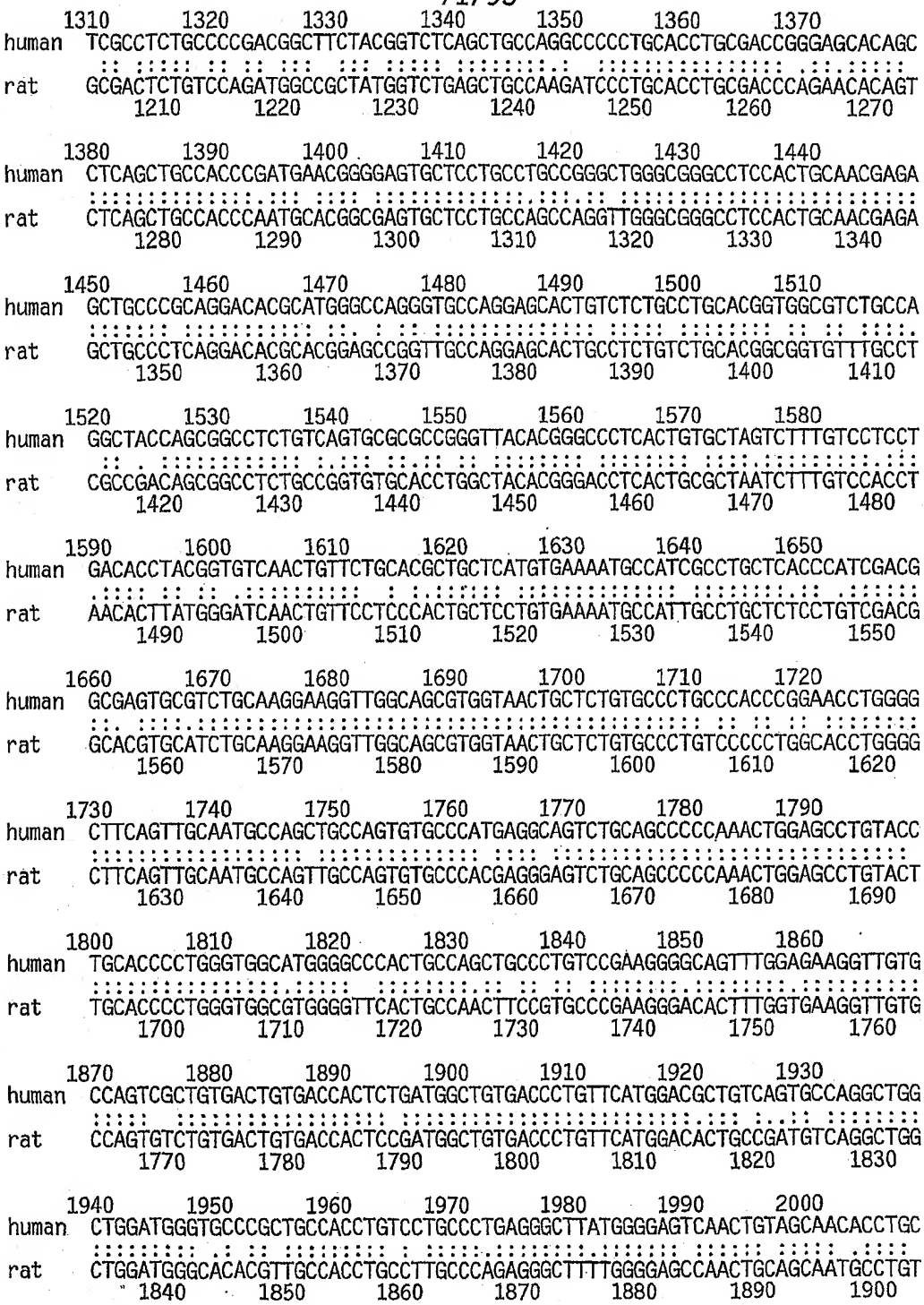


FIG.34C

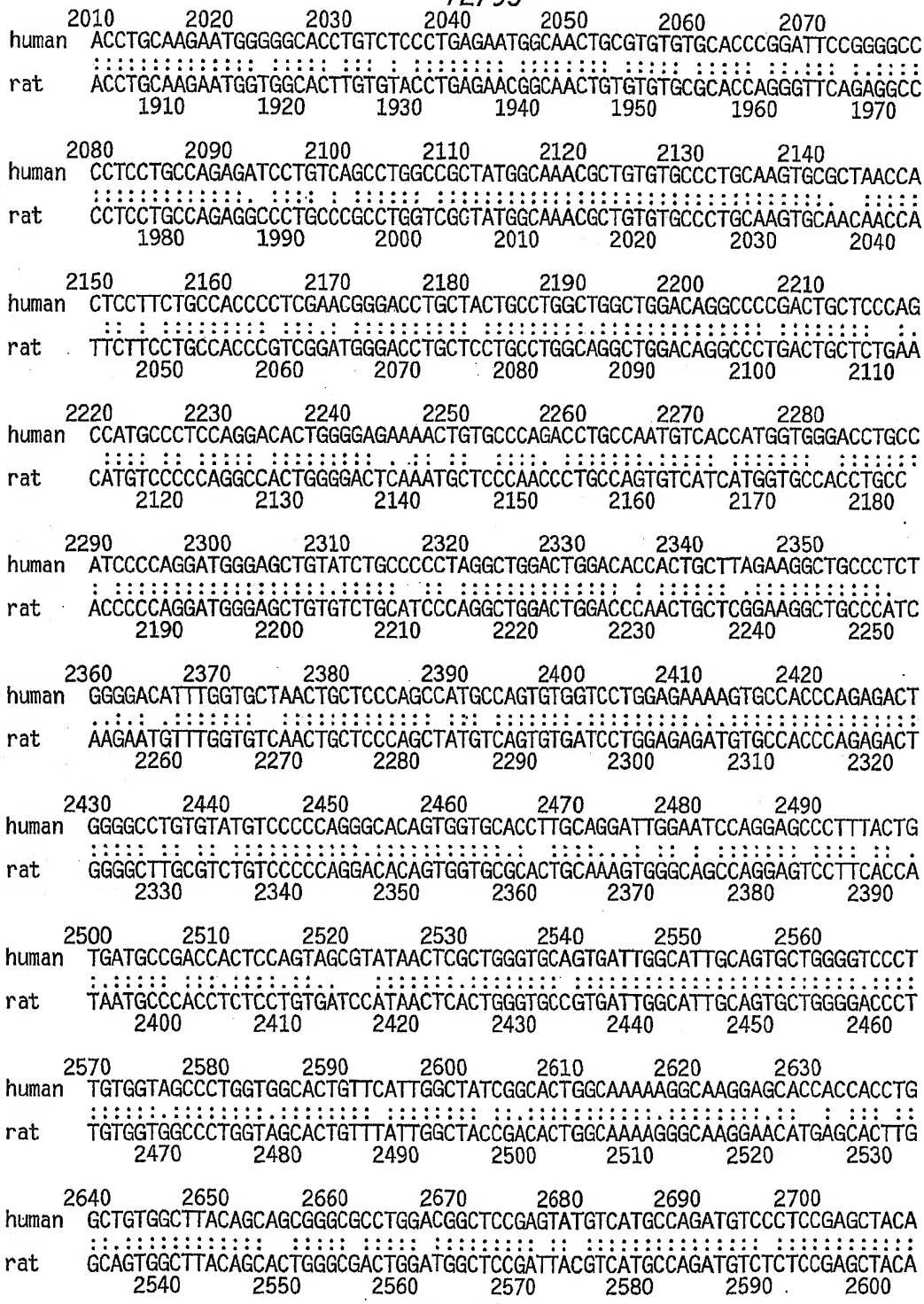


FIG. 34D

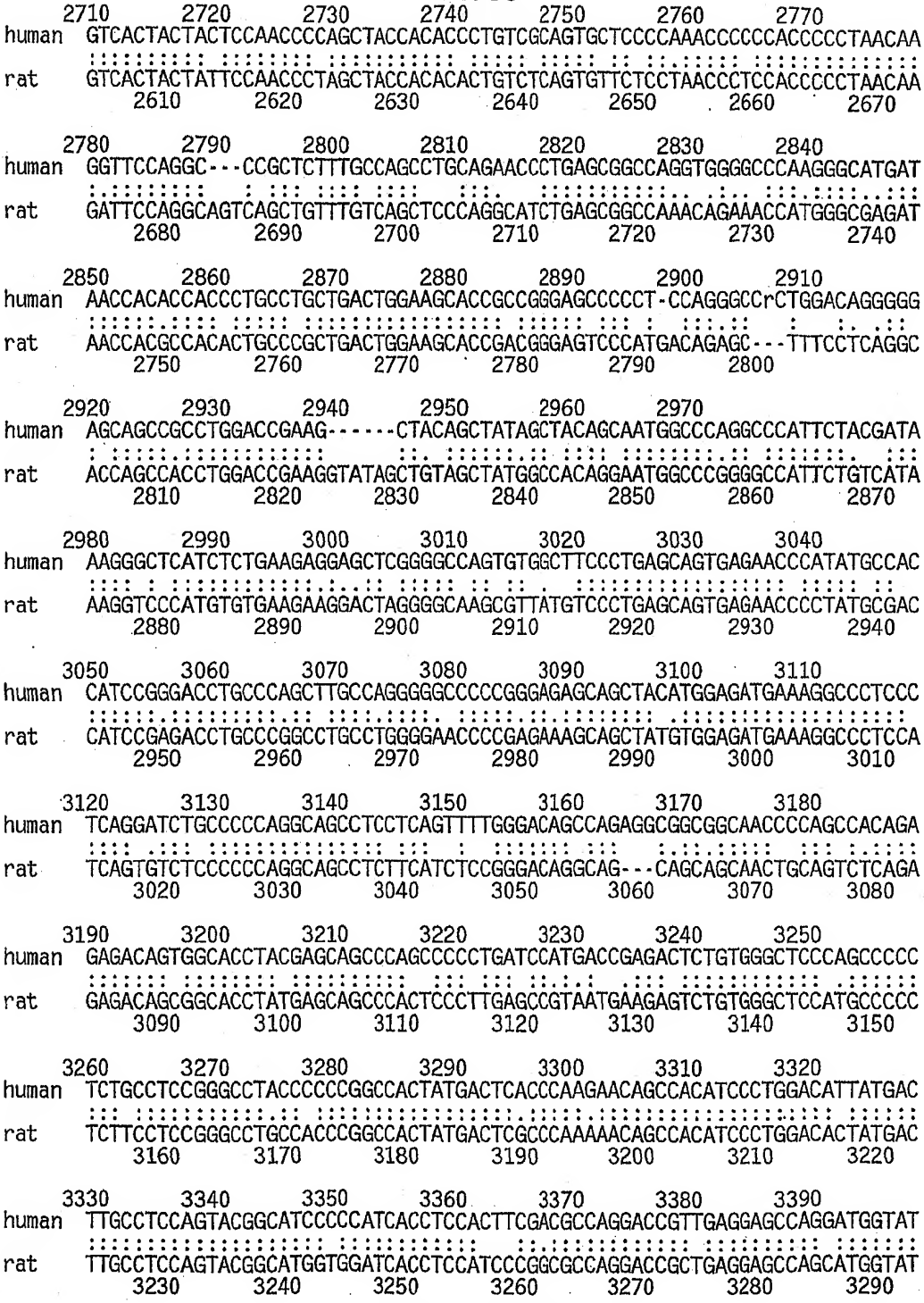


FIG.34E

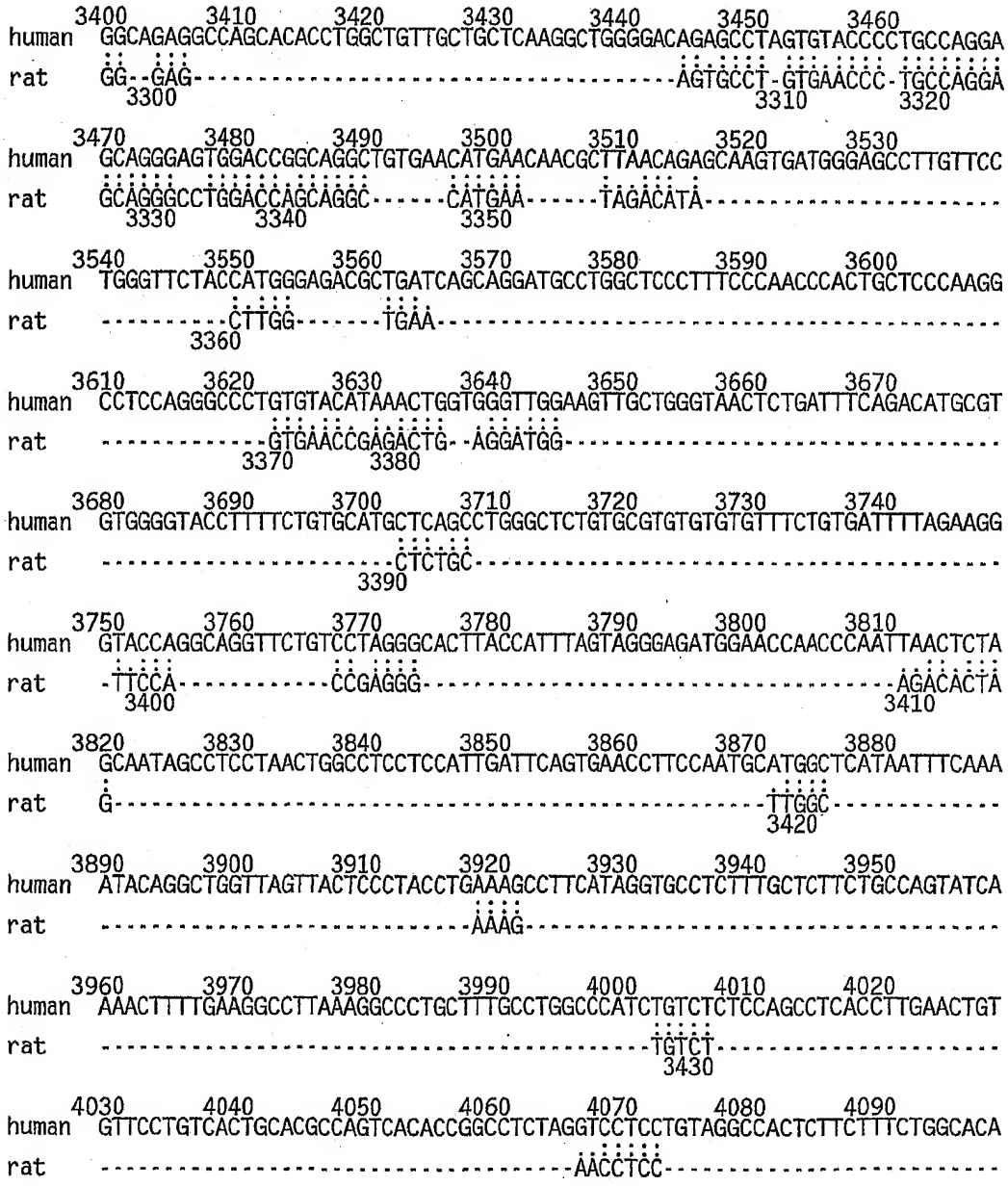


FIG.34F

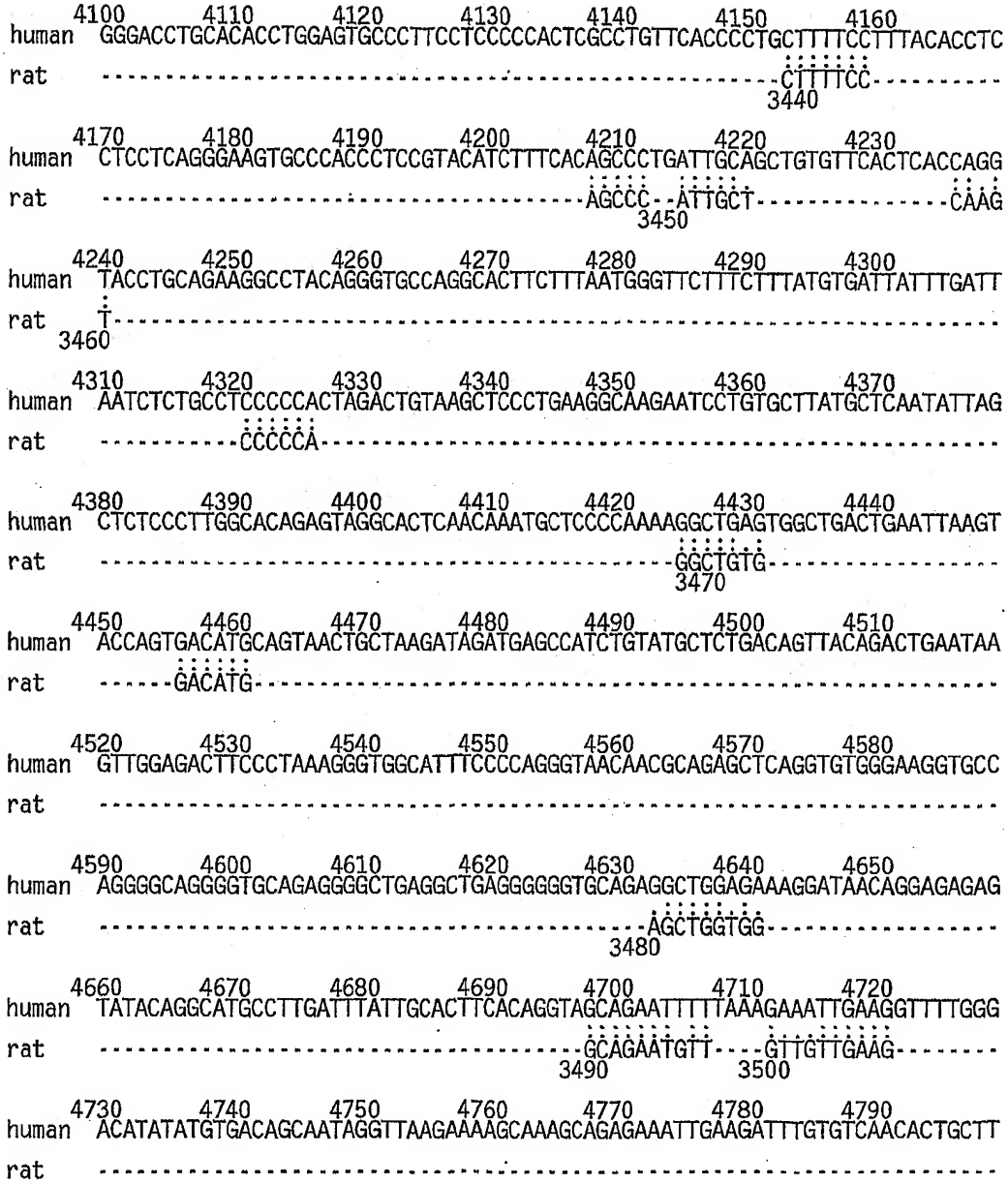


FIG.34G

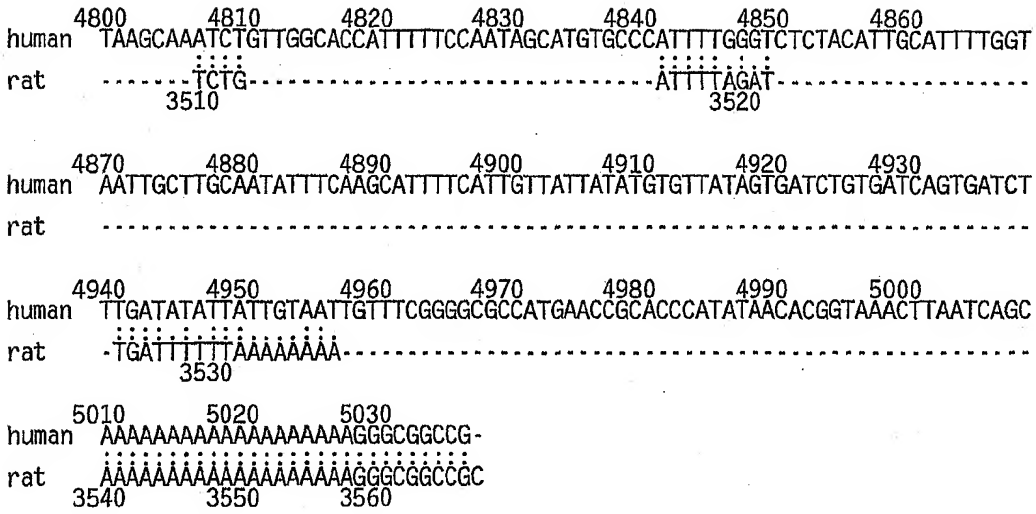

4800 4810 4820 4830 4840 4850 4860
human TAAGCAAATCTGTTGGCACCATTTTTCCAATAGCATGTGCCCATTTTGGGTCTCTACATTGCATTTTGGT
rat -------TCTG----------ATTTTAGAT----------
3510 3520
4870 4880 4890 4900 4910 4920 4930
human AATTGCTTGCAATATTTCAAGCATTTTCATTGTTATTATATGTGTTATAGTGATCTGTGATCAGTGATCT
rat ----------
4940 4950 4960 4970 4980 4990 5000
human TTGATATATTATTGTAATTGTTTCGGGGCGCCATGAACCGCACCCATATAACACGGTAAACTTAATCAGC
rat -TGATTTTTTAAAAAAAA----------
3530
5010 5020 5030
human AAAAAAAAAAAAAAAAAAAAGGGCGGCCG-
rat AAAAAAAAAAAAAAAAAAAAGGGCGGCCGC
3540 3550 3560

FIG.34H

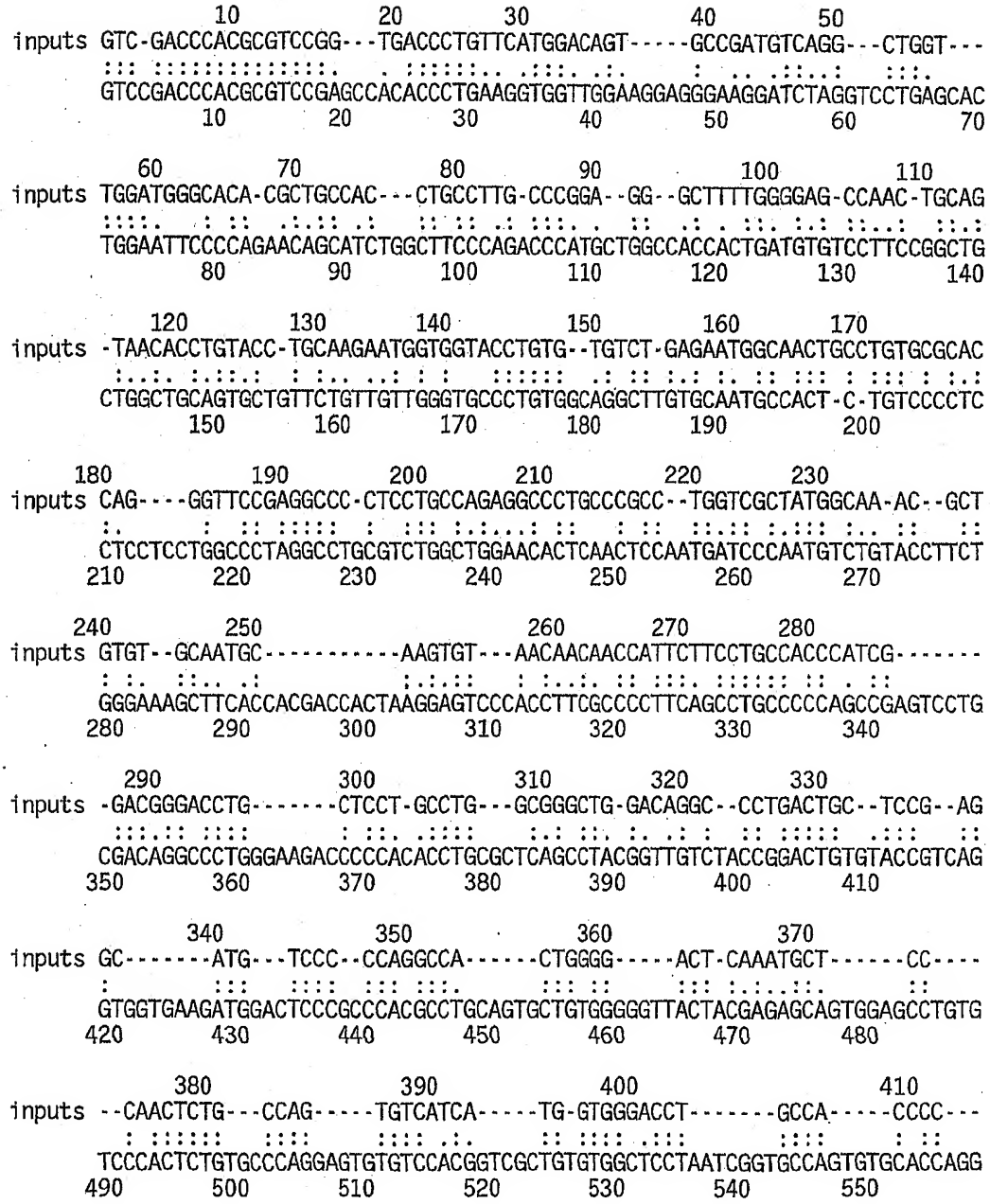


FIG.35A

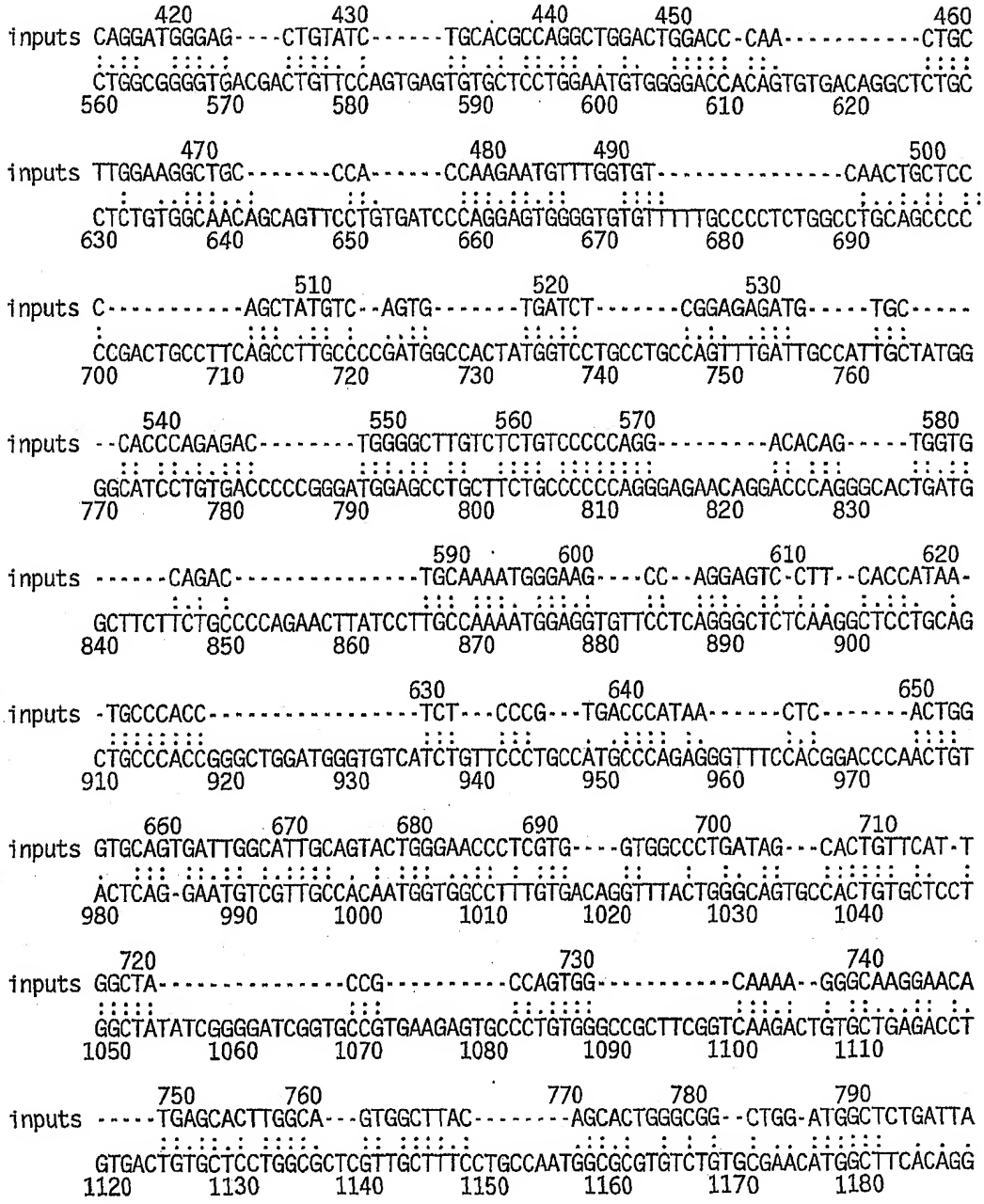


FIG.35B

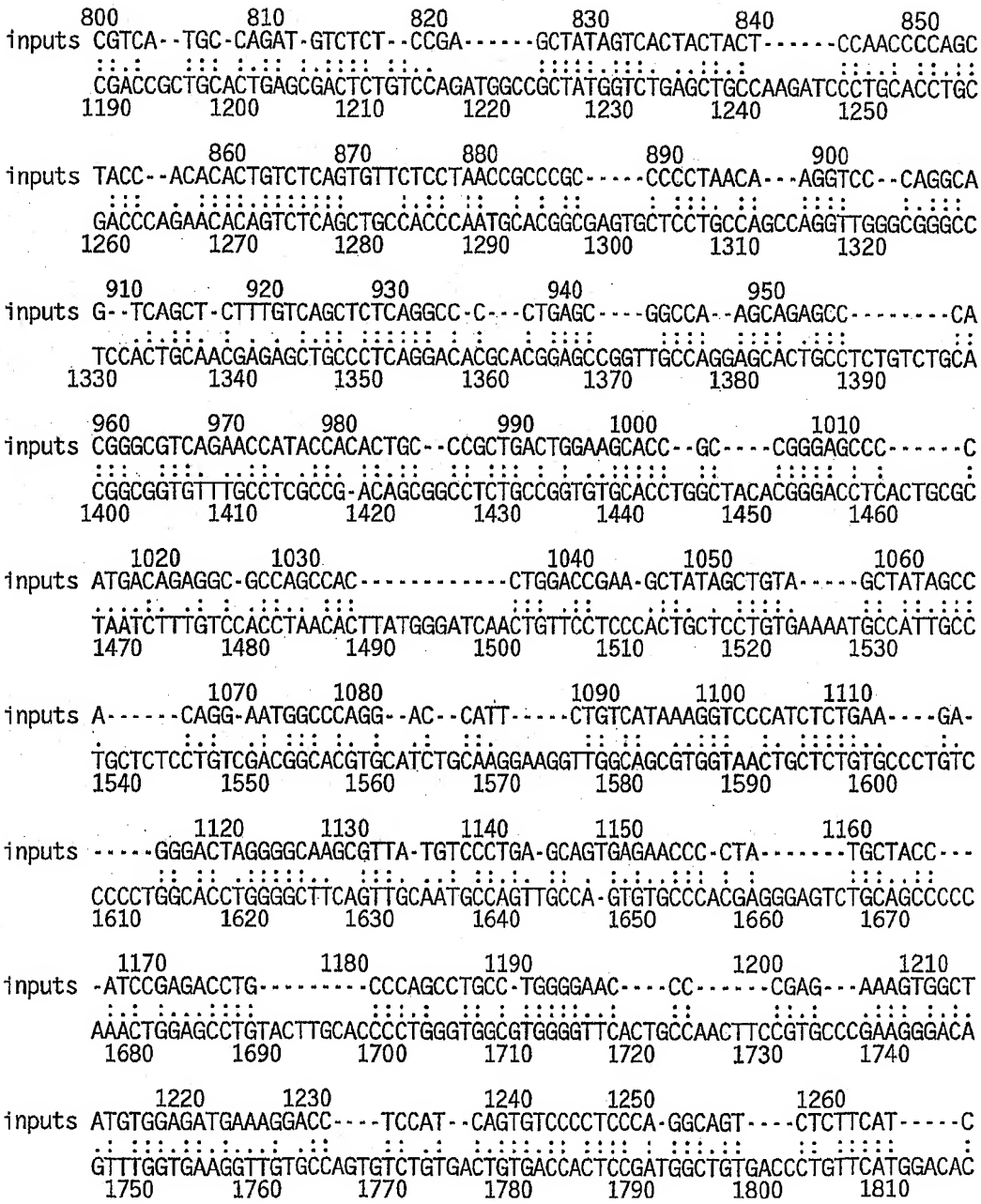


FIG.35C

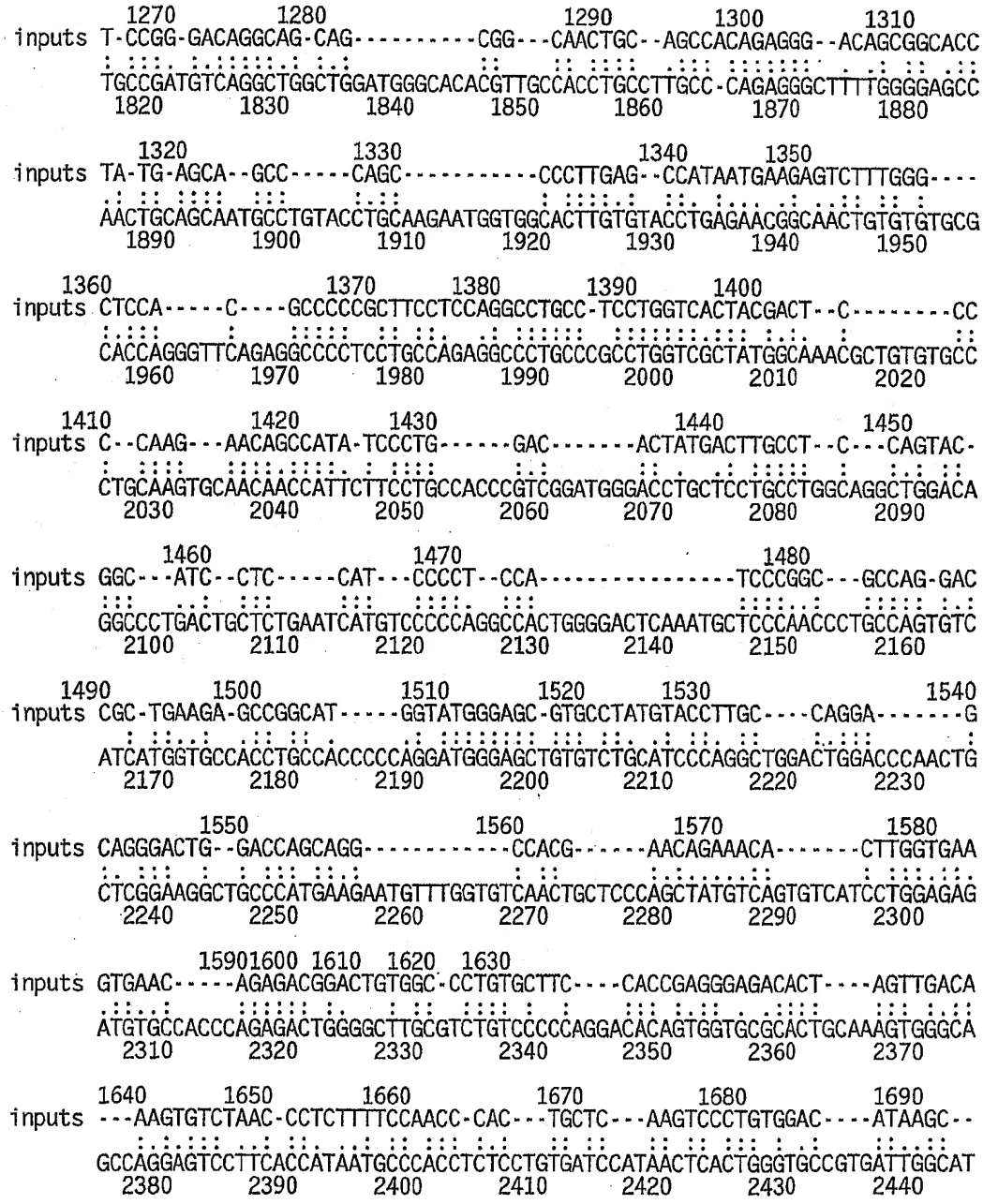

1270 1280 1290 1300 1310
inputs T-CCGG-GACAGGCAG-CAG----------CGG---CAACTGC--AGCCACAGAGGG--ACAGCGGCACC
TGCCGATGTCAGGCTGGCTGGATGGGCACACGTTGCCACCTGCCTTGCC-CAGAGGGCTTTTGGGGAGCC
1820 1830 1840 1850 1860 1870 1880
1320 1330 1340 1350
inputs TA-TG-AGCA--GCC-----CAGC----------CCCTTGAG--CCATAATGAAGAGTCTTTGGG----
AACTGCAGCAATGCCTGTACCTGCAAGAATGGTGGCACTTGTGTACCTGAGAACGGCAACTGTGTGTGCG
1890 1900 1910 1920 1930 1940 1950
1360 1370 1380 1390 1400
inputs CTCCA-----C----GCCCCCGCTTCCTCCAGGCCTGCC-TCCTGGTCACTACGACT--C--------CC
CACCAGGGTTCAGAGGCCCCTCCTGCCAGAGGCCCTGCCCGCCTGGTCGCTATGGCAAACGCTGTGTGCC
1960 1970 1980 1990 2000 2010 2020
1410 1420 1430 1440 1450
inputs C--CAAG---AACAGCCATA-TCCCTG------GAC-------ACTATGACTTGCCT--C---CAGTAC-
CTGCAAGTGCAACAACCATTCTTCCTGCCACCCGTCGGATGGGACCTGCTCCTGCCTGGCAGGCTGGACA
2030 2040 2050 2060 2070 2080 2090
1460 1470 1480
inputs GGC---ATC--CTC-----CAT---CCCCT--CCA---------------TCCCGGC---GCCAG-GAC
GGCCCTGACTGCTCTGAATCATGTCCCCCAGGCCACTGGGGACTCAAATGCTCCCAACCCTGCCAGTGTC
2100 2110 2120 2130 2140 2150 2160
1490 1500 1510 1520 1530 1540
inputs CGC-TGAAGA-GCCGGCAT-----GGTATGGGAGC-GTGCCTATGTACCTTGC----CAGGA-------G
ATCATGGTGCCACCTGCCACCCCCAGGATGGGAGCTGTGTCTGCATCCCAGGCTGGACTGGACCCAACTG
2170 2180 2190 2200 2210 2220 2230
1550 1560 1570 1580
inputs CAGGGACTG--GACCAGCAGG-----------CCACG------AACAGAAACA-------CTTGGTGAA
CTCGGAAGGCTGCCCATGAAGAATGTTTGGTGTCAACTGCTCCCAGCTATGTCAGTGTCATCCTGGAGAG
2240 2250 2260 2270 2280 2290 2300
15901600 1610 1620 1630
inputs GTGAAC-----AGAGACGGACTGTGGC-CCTGTGCTTC----CACCGAGGGAGACACT----AGTTGACA
ATGTGCCACCCAGAGACTGGGGCTTGCGTCTGTCCCCCAGGACACAGTGGTGCGCACTGCAAAGTGGGCA
2310 2320 2330 2340 2350 2360 2370
1640 1650 1660 1670 1680 1690
inputs ---AAGTGTCTAAC-CCTCTTTTCCAACC-CAC---TGCTC---AAGTCCCTGTGGAC----ATAAGC--
GCCAGGAGTCCTTCACCATAATGCCCACCTCTCCTGTGATCCATAACTCACTGGGTGCCGTGATTGGCAT
2380 2390 2400 2410 2420 2430 2440

FIG.35D

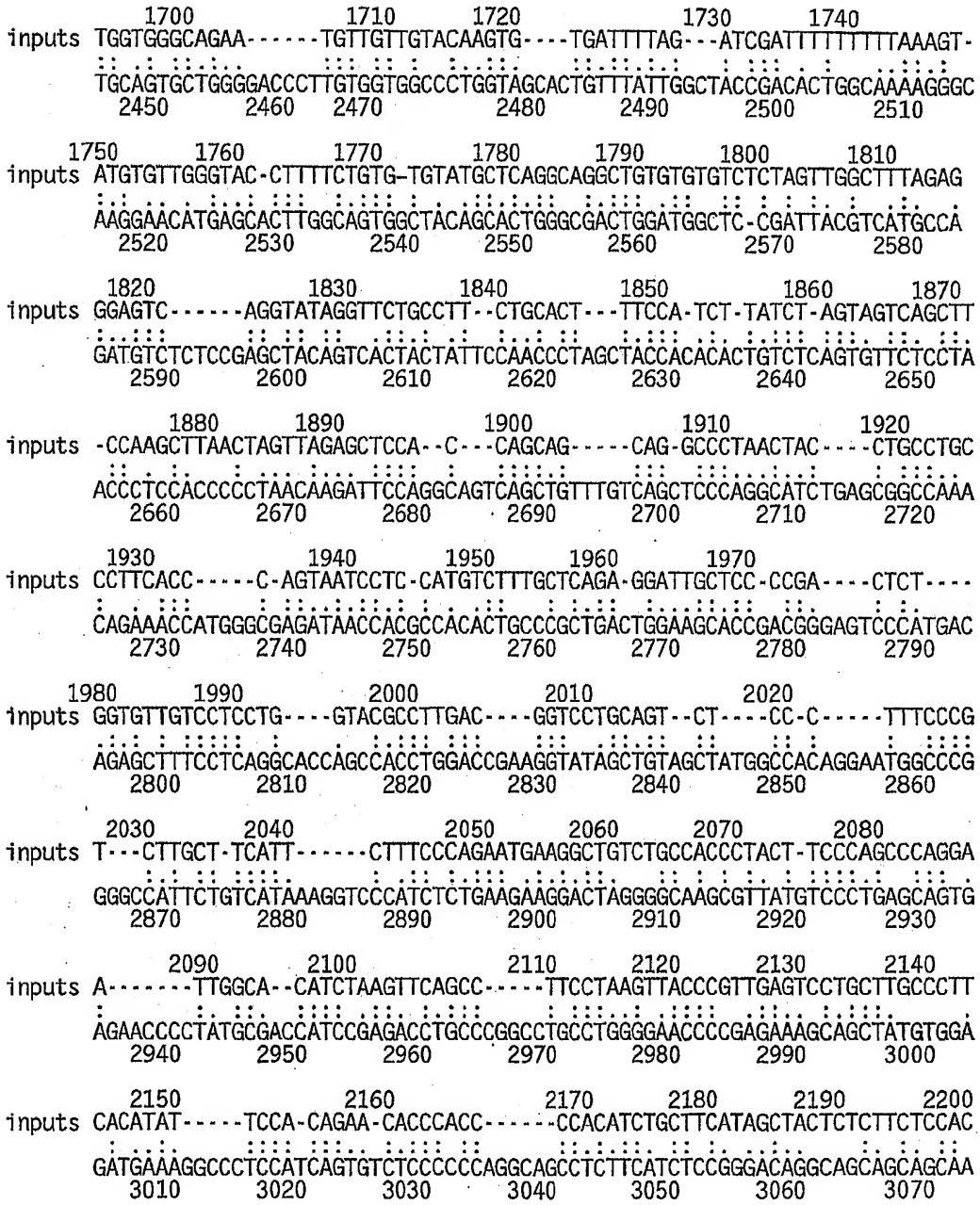


FIG.35E

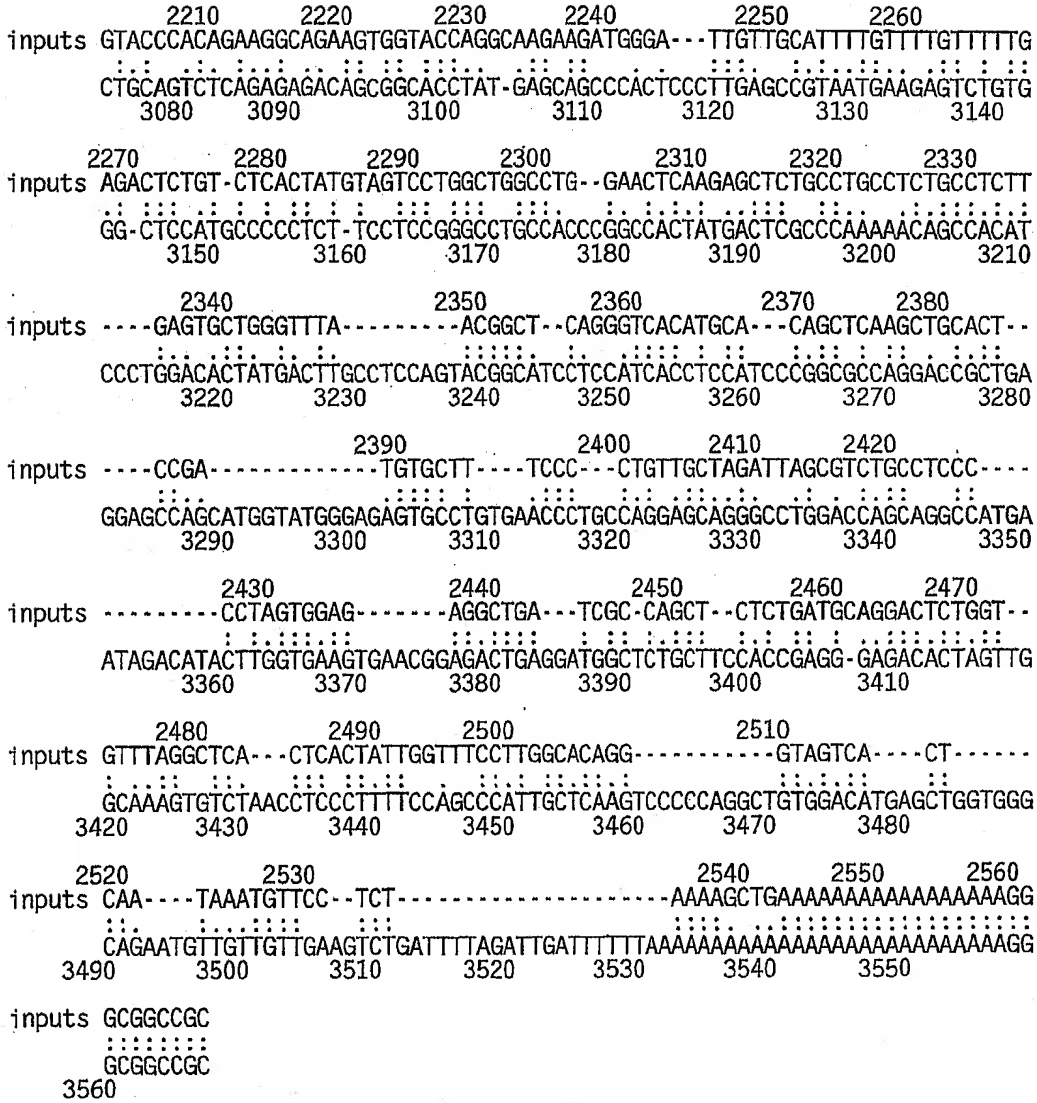

inputs GTACCCACAGAAGGCAGAAGTGGTACCAGGCAAGAAGATGGGA---TTGTTGCATTTTGTTTTGTTTTTG
CTGCAGTCTCAGAGAGACAGCGGCACCTAT-GAGCAGCCCACTCCCTTGAGCCGTAATGAAGAGTCTGTG
inputs AGACTCTGT-CTCACTATGTAGTCCTGGCTGGCCTG--GAACTCAAGAGCTCTGCCTGCCTCTGCCTCTT
GG-CTCCATGCCCCCTCT-TCCTCCGGGCCTGCCACCCGGCCACTATGACTCGCCCAAAAACAGCCACAT
inputs ----GAGTGCTGGGTTTA---------ACGGCT--CAGGGTCACATGCA---CAGCTCAAGCTGCACT--
CCCTGGACACTATGACTTGCCTCCAGTACGGCATCCTCCATCACCTCCATCCCGGCGCCAGGACCGCTGA
inputs ----CCGA------------TGTGCTT----TCCC---CTGTTGCTAGATTAGCGTCTGCCTCCC----
GGAGCCAGCATGGTATGGGAGAGTGCCTGTGAACCCTGCCAGGAGCAGGGCCTGGACCAGCAGGCCATGA
inputs ---------CCTAGTGGAG-------AGGCTGA---TCGC-CAGCT--CTCTGATGCAGGACTCTGGT--
ATAGACATACTTGGTGAAGTGAACGGAGACTGAGGATGGCTCTGCTTCCACCGAGG-GAGACACTAGTTG
inputs GTTTAGGCTCA---CTCACTATTGGTTTCCTTGGCACAGG----------GTAGTCA----CT------
GCAAAGTGTCTAACCTCCCTTTTCCAGCCCATTGCTCAAGTCCCCCAGGCTGTGGACATGAGCTGGTGGG
inputs CAA----TAAATGTTCC--TCT-------------------AAAAGCTGAAAAAAAAAAAAAAAAAAGG
CAGAATGTTGTTGTTGAAGTCTGATTTTAGATTGATTTTTTAAAAAAAAAAAAAAAAAAAAAAAAAAGG
inputs GCGGCCGC
GCGGCCGC

FIG.35F

```
                10        20        30        40        50        60        70
inputs MAPARAGFCPLLLLLLLGLWVAEIPVSAKPKGMTSSQWFKIQHMQPSPQACNSAMKNINKHTKRCKDLNT
       :.     .:  :::::: :: .  :. : :: .: ..:: :::.:::::  :: :: .::. ...::. ::
       MV-----LCFPLLLLLLVLWGPVCPLHAWPKRLTKAHWFEIQHIQPSPLQCNRAMSGINNYAQHCKHQNT
                    10        20        30        40        50        60

                80        90       100       110       120       130       140
inputs FLHEPFSSVAATCQTPKIACKNGDKNCHQSHGPVSLTMCKLTSGKYPNCRYKEKRQNKSYVVACKPPQKK
       :::..: .:::.:.   :.:::   ::::: ::..: :.::::::::.::: .  : : ..::: ::::
       FLHDSFQNVAAVCDLLSIVCKNRRHNCHQSSKPVNMTDCRLTSGKYPQCRYSAAAQYKFFIVACDPPQKS
          70        80        90       100       110       120       130

               150
inputs DSQQFHLVPVHLDRVL
       :.. . ::::::: .:
       DPP-YKLMVPVHLDSIL
         140       150
```

FIG.36

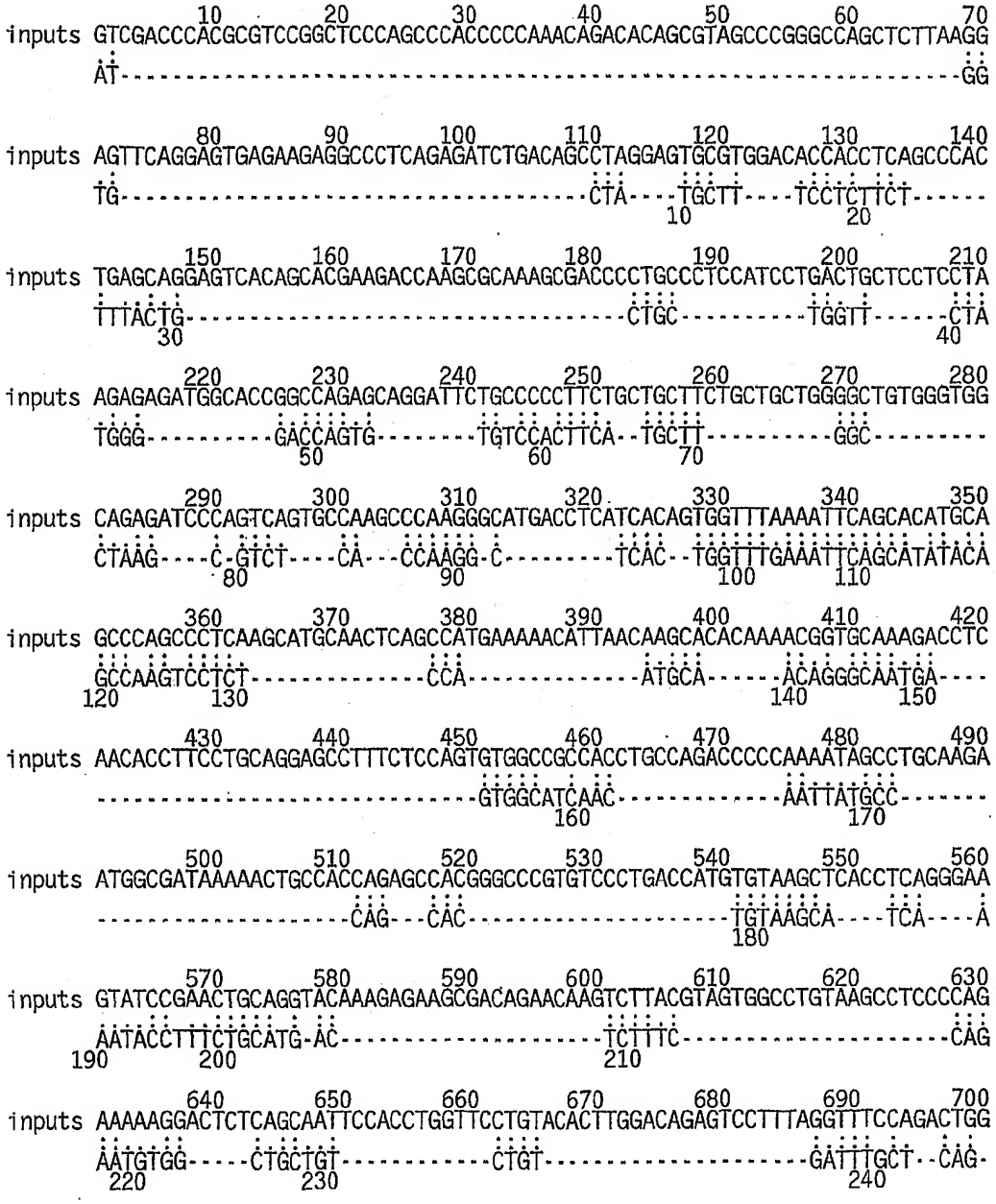


FIG.37A

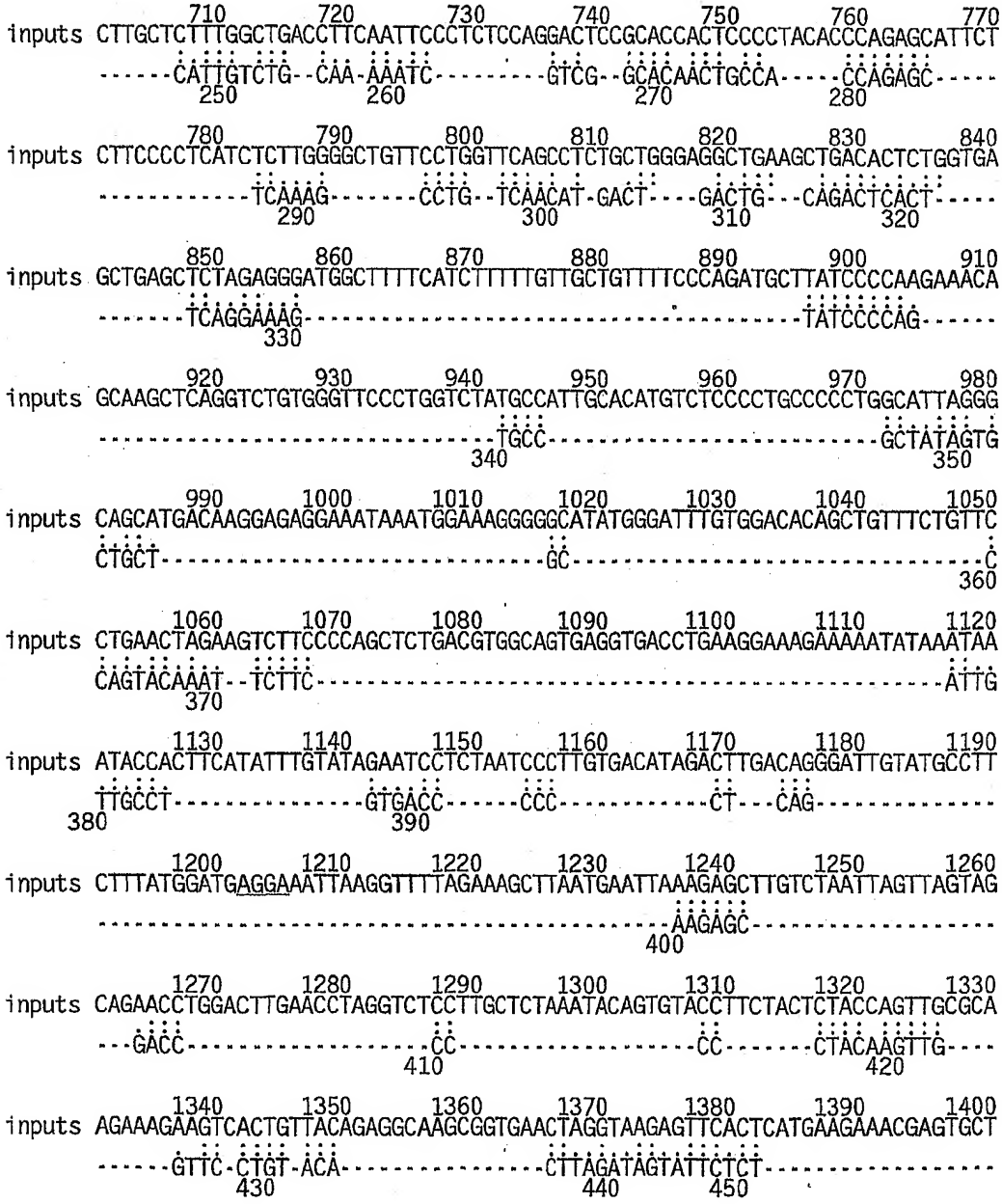


FIG.37B

```
                1410        1420        1430        1440        1450        1460        1470
inputs CTGAAGAGCCAGTTACCCTGTGTTGGCTGCAATAAAGGTCATTACCTCTCTAGCCAAAAAAAAAAAAAAA
       ------------------------------------------------------------------------

                1480        1490
inputs AAAAAAAAAAAAAAAAAAAAAAAAAA
                                ::
       ------------------------AA
```

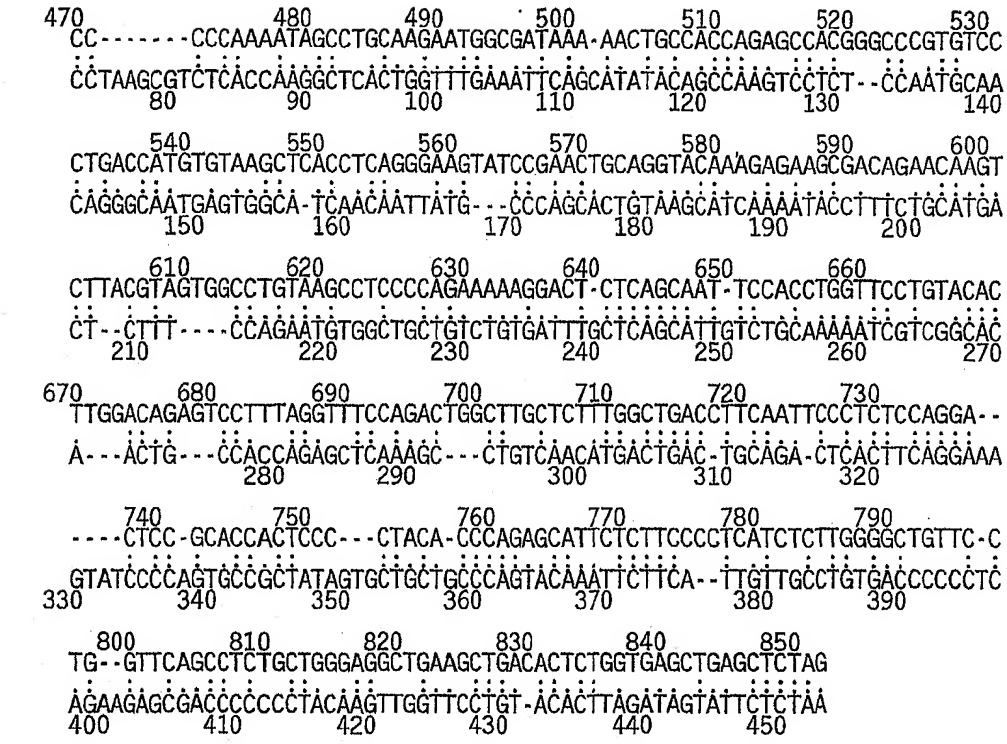
FIG.37C

```
     240        250        260        270        280        290        300
AGGATTCTGCCCCCTTCTGCTGCTTCTGCTGCTGGGGCTGTGGGTGGCAGAGATCCCAGTCAGTGCCAAG
.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.
GGTGCTATGCTTTCCTCTTCTTTTACTGCTGCTGGTTCTATGGGGACCAGTGTGTCCACTTCATGCTTGG
      10         20         30         40         50         60         70

     310        320        330        340        350        360        370
CCCAAGGGCATGACCTCATCACAGTGGTTTAAAATTCAGCACATGCAGCCCAGCCCTCAAGCATGCAACT
.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.
CCTAAGCGTCTCACCAAGGCTCACTGGTTTGAAATTCAGCATATACAGCCAAGTCCTCTCCAATGCAACA
      80         90        100        110        120        130        140

     380        390        400        410        420        430        440
CAGCCATGAAAAACATTAACAAGCACACAAAACGGTGCAAAGACCTCAACACCTTCCTGCACGAGCCTTT
.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.
GGGCAATGAGTGGCATCAACAATTATGCCCAGCACTGTAAGCATCAAAATACCTTTCTGCATGACTCTTT
      150        160        170        180        190        200        210

     450        460        470        480        490        500        510
CTCCAGTGTGGCCGCCACCTGCCAGACCCCCAAAATAGCCTGCAAGAAT-GGCGATAAAAACTGCCACCA
.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.
CCAGAATGTGGCTGCTGTCTGTGATTTGCTCAGCATTGTCTGCAAAAATCGTCGGCACAA-CTGCCACCA
      220        230        240        250        260        270        280

     520        530        540        550        560        570        580
GAGCCACGGGCCCGTGTCCCTGACCATGTGTAAGCTCACCTCAGGGAAGTATCCGAACTGCAGGTACAAA
.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.
GAGCTCAAAGCCTGTCAACATGACTGACTGCAGACTCACTTCAGGAAAGTATCCCCAGTGCCGCTATAGT
      290        300        310        320        330        340        350

      590        600        610        620        630        640        650
G-AGAAGCGACAGAACAAGTCTTACGTAGTGGCCTGTAAGCCTCCCCAGAAAAAGGACTCTCAGCAATTC
.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.
GCTGCTGC-CCAGTACAAATTCTTCATTGTTGCCTGTGACCCCCCTCAGAAGAGCGACCCCCC-C--TAC
       360        370        380        390        400        410

      660        670        680
CACCTGGTTCCTGTACACTTGGACAGAGTCCTTTAG
.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.:.
AAGTTGGTTCCTGTACACTTAGATAGTATTCTCTAA
420        430        440        450
```

43.4% identity in 477 aa overlap; score: 746

```
     410        420        430        440        450        460
GGTGCAAAG---ACCTCAACACCTTC--CTGCACGAGCCTTTC--TCCAGTGTGGCCGCCACCTGCCAGA
.:.:.:.:.---:.:.:.:.:.:.:--:.:.:.:.:.:.:--:.:.:.:.:.:.:.:.:.:.:.:.:.:.
GGTGCTATGCTTTCCTCTTCTTTTACTGCTGCTGGTTCTATGGGGACCAGTGTGTCCACTTCATGCTTGG
       10         20         30         40         50         60         70
```

FIG.38A



46.5% identity in 488 aa overlap; score: 709

FIG.38B

```
laminin_EGF: domain 1 of 4, from 3 to 37: score -1.2, E = 0.59

                  *->CdCnphGslsddtCdsddelfgeetGqClkCkpnvtGrrCdr.CkpG
                       +  G       d+       ++GqC+ C+ + +G+rC  +C +G
      mT272       3    ---HASG-------DP-------VHGQCR-CQAGWMGTRCHLpCPEG  31

                  yyglpsgdpgqgC<-*
                  ++g      + +C
      mT272      32 FWG------A-NC    37


EGF: domain 1 of 4, from 37 to 67: score 19.2, E = 0.1

                  *->CapnnpCsngGtCvntpggssdnfggytCeCppGdyylsytGkrC<-
                     C+ ++ C+ngGtCv+  g           C+C+pG      + G+ C
      mT272      37    CSNTCTCKNGGTCVSENG---------NCVCAPG-----FRGPSC    67


DSL: domain 1 of 1, from 10 to 67: score -21.2, E = 8.1

                  *->WstdkhiggrtslGfnleyrirvtCdenYYGsgCnkFCrPrdDafgH
                        +   + +      + r +  C e   G+ C++ C       +g+
      mT272      10    --HGQCRCQAG----WMGTRCHLPCPEGFWGANCSNTCTCK---NGG  47

                  ytCdenGnklCleGWkGeyC<-*
                     +enGn  C++G +G+ C
      mT272      48 TCVSENGNCVCAPGFRGPSC    67


laminin_EGF: domain 2 of 4, from 41 to 80: score -1.5, E = 0.63

                  *->CdCnphGslsddtCdsddelfgeetGqClkCkpnvtGrrCdr.CkpG
                     C+C + G     tC s       e G C+ C p++ G+ C r+C pG
      mT272      41    CTCKNGG-----TCVS-------ENGNCV-CAPGFRGPSCQRpCPPG  74

                  yyglpsgdpgqgC<-*
                   y     + +  C
      mT272      75 RY-----GKR--C    80


EGF: domain 2 of 4, from 80 to 110: score 11.8, E = 1.9

                  *->CapnnpCsng.GtCvntpggssdnfggytCeCppGdyylsytGkrC<
                     C +   C+n++  C+++ g          Tc C  G      +tG++C
      MT272      80    CVQC-KCNNNhSSCHPSDG---------TCSCLAG-----WTGPDC   110
```

FIG.39A

```
laminin_EGF: domain 3 of 4, from 83 to 123: score 25.6, E = 0.0012

                   *->CdCnphGalsddtCdsddelfgeetGqClkCkpnvtGrrC.drCkpG
                      C Cn++     ++C++       +G C+ C+ + tG++C++ C pG
      mT272     83     CKCNNNH----SSCHP-------SDGTCS-CLAGWTGPDCsEACPPG  117

                    yyglpsgdpgqgC<-*
                    ++gl        C
      mT272    118   HWGL-------KC     123


EGF: domain 3 of 4, from 123 to 153: score 27.3, E = 0.00036

                   *->CapnnpCsngGtCvntpggssdnfggytCeCppGdyylsytGkrC<-
                      C++++ C++gGtc++  g         +C+C+pG     +tG++c
      mT272     123    CSQLCQCHHGGTCHPQDG---------SCICTPG-----WTGPNC   153


laminin_EGF: domain 4 of 4, from 127 to 172: score -5.5, E = 1.4

                    *->CdCnphGslsddtCdsddelfgeetGqClkCkpnvtGrrC.drCkpG
                       C+C++ G          tC++        G C  C p+ tG++C + C p

      mT272     127     CQCHHGG-----TCHP-------QDGSCI-CTPGWTGPNClEGCPPR 160

                    yyglpsg.dpgqgC<-*
                     +g  ++++ + +C
      mT272     161   MFG-VNCsQLC-QC    172


EGF: domain 4 of 4, from 166 to 196: score 4.5, E = 5.8

                   *->CapnnpCsngGtCvntpggssdnfggytCeCppGdyylsytGkrC<-
                      C++++ C+ g  C++  g          C+CppG      +G +C
      mT272     166    CSQLCQCDLGEMCHPETG---------ACVCPPG-----HSGADC     196
```

FIG.39B

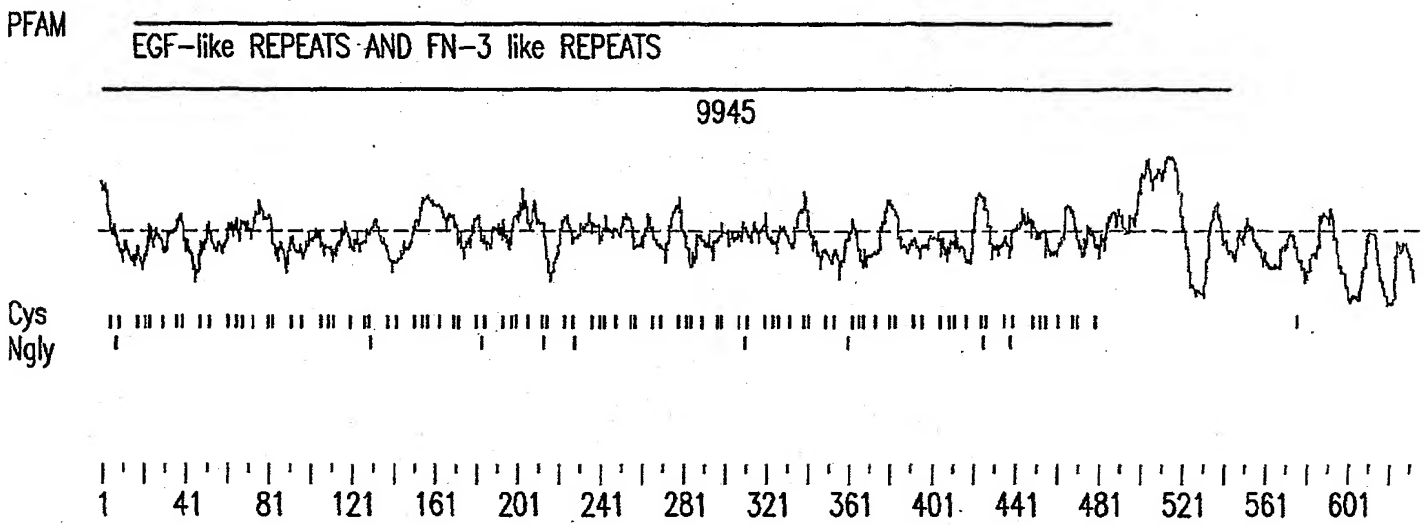


FIG.40

```
                    *->CaPnnpCsnqGtCvntpggssdntggytCeCppGayylsytGkrC<-
                       C++++ C+ngG C    g          +C+C+pG     y+G+rC
ratT272     18        !ECRCHNGGLCDRFTG---------QCHCAPG-----YIGDPRC     48


laminin_EGF: domain 1 of 11, from 22 to 61: score 12.3, E = 0.038

                    *->CdCnphGslsddtCdsddelfgeetGqClkCkpnvtGrrC.drCkpG
                       C C++ G      Cd+       +tGqC+ C p++ G+rC+++C  G
ratT272     22        CRCHNGG-----LCDR-------FTGQCH-CAPGYIGDRCrEECPVG   55

                    yyglpsgdpgqgC<-*
                     +g       q+C
ratT272     56 RFG-------QDC    61


EGF: domain 2 of 11, from 61 to 91: score 18.3, E = 0.18

                    *->CapnnpCsngGtCvntpggssdnfggytCeCppGdyylsytGkrC<-
                       Ca+++ C  q++C +  q          C C +G      +tG+rC
ratT272     61        CAETCDCAPGARCFPANG---------ACLCEHG-----FTGDRC     91


laminin_EGF: domain 2 of 11, from 65 to 105: score 4.0, E = 0.2

                    *->CdCnphGslsddtCdsddelfgeetGqClkCkpnvtGrrCdr..Ckp
                       CdC p +     +C +         G+Cl C +++tG+rC ++ C +
ratT272     65        CDCAPGA-----RCFP-------ANGACL-CEHGFTGDRCTErlCPD   98

                    Gyyglpsgdpgqgc<-*
                    G ygl          +C
ratT272     99 GRYGL-------SC   105


EGF: domain 3 of 11, from 105 to 137: score 4.1, E = 9.6

                    *->CapnnpCsng..GtCvntpggssdnfggytCeCppGdyylsytGkrC
                       C++++ C+  ++  C++ +g          +C C+pG      ++G +C
ratT272    105      CQDPCTCDPEhsLSCHPMHG---------ECSCQPG-----WAGLHC  137


laminin_EGF: domain 3 of 11, from 109 to 150: score 13.1, E = 0.032

                    *->CdCnphGslsddtCdsddelfgeetGqClkCkpnvtGrrCdr.CkpG
                       C+C+p  sls   C++        ++G+C+ C+p ++G +C+++C
ratT272    109      CTCDPEHSLS---CHP-------MHGECS-CQPGWAGLHCNEsCP--  142

                    yyglpsgdpqqgC<-*
                      ++   + g gC
ratT272    143 --QD---THGAGC    150
```

FIG.41A

```
EGF: domain 4 of 11, from 150 to 180: score 27.7, E = 0.00026

                   *->CapnnpCsngGtCvntpggssdnfggytCeCppGdyylsytGkrC<-
                      C++++ C++gG+C+   g          C+C+pG     ytG++C
     ratT272   150    CQEHCLCLHGGVCLADSG---------LCRCAPG-----YTGPHC   180


laminin_EGF: domain 4 of 11, from 154 to 193: score 8.4, E = 0084

                   *->CdCnphGslsddtCdsddelfgeetGqClkCkpnvtGrrC.drCkpG
                      C C +hg    + C          +G C+ C p++tG++C + C p+
     ratT272   154    CLC-LHG----GVCLA-------DSGLCR-CAPGYTGPHaNLCPPN   187

                   yyglpsgdpgqgC<-*
                   *yg        +C
     ratT272   188 TYGI-------NC    193


EGF: domain 5 of 11, from 193 to 223: score 10.6, E = 2.5

                   *->CapnnpCsngGtCvntpggssdnfggytCeCppGdyylsytGkrC<-
                      C++++ C n   C ++ g         tC+C++G     ++  +C
     ratT272   193    CSSHCSCENAIACSPVDG---------TCICKEG-----WQRGNC   223


laminin_EGF: domain 5 of 11, from 197 to 236: score 0.7, E = 0.4

                   *->CdCnphGslsddtCdsddelfgeetGqClkCkpnvtGrrCdr.CkpG
                      C C ++       C +       + G C  Ck++ +  +C  +C pG
     ratT272   197    CSCENAI-----ACSP-------VDGTCI-CKEGWQRGNCSVpCPPG 230

                   yyglpsgdpgqgC<-*
                   ++g+       +C
     ratT272   231 TWG--------SC    236


EGF: domain 6 of 11, from 236 to 266: score 11.8, E = 1.9

                   *->CapnnpCsngGtCvntpggssdnfggytCeCppGdyylsytGkrC<-
                      C+  + C + G+C +  g          C+C+pG     + G +C
     ratT272   236    CNASCQCAHEGVCSPQTG---------ACTCTPG-----WRGVHC   266


laminin_EGF: domain 6 of 11, from 240 to 279: score -2.2, E = 0.73

                   *->CdCnphGslsddtCdsddelfgeetGqClkCkpnvtGrrCdr.CkpG
                      C+C + G      C +        tG+C  C p+  G +C  +C  G
     ratT272   240    CQCAHEG-----VCSP-------QTGACT-CTPGWRGVHCQLpCPKG 273

                   yyglpsgdpgqgC<-*
                    +g        +gC
     ratT272   274 QFG-------EGC    279
```

FIG.41B

```
DSL: domain 1 of 1, from 246 to 309: score -19.4, E = 5.2

                   *->WstdkhiggrtslGfnleyrirvtCdenYYGegCnkFCrPrdDafgH
                      +  ++++g+ t       +++    C +  +GegC+  C+       H
     ratT272   246    GVCSPQTGACTCTPGQRGVHCQLPCPKGQFGEGCASVCDCD-----H 287

                   yt.Cd.enGnklcleGWkGeyC<-*
                   + +Cd+ +G  +C +GW+G  C
     ratT272   288 SDgCDpVHGHCRCQAGWMGTRC    309


EGF: domain 7 of 11, from 279 to 309: score 7.0, E = 5.3

                   *->CapnnpCsngGtCvntpggssdnfggytCeCppGdyylsytGkrC<-
                      Ca+ + C++   C +++g         +C+C+ G     + G rC
     ratT272   279    CASVCDCDHSDGCDPVHG---------HCRCQAG-----WMGTRC   309


laminin_EGF: domain 7 of 11, from 283 to 322: score 12.7, E = 0.035

                   *->CdCnphGslsddtCdsddelfgeetGqClkCkpnvtGrrCdr.CkpG
                      CdC+ h+    d Cd+       ++G+C+ C+ + +G+rC  +C +G
     ratT272   283    CDCD-HS----DGCDP-------VHGHCR-CQAGQMGTRCHLpCPEG 316

                   yyglpsgdpgqgC<-*
                   ++g      + +C
     ratT272   317 FWG------A-NC    322


EGF: domain 8 of 11, from 322 to 352: score 17.3, E = 0.38

                   *->CapnnpCsngGtCvntpggssdnfggytCeCppGdyylsytGkrC<-
                      C+  + C+ngGtCv+  g           C+C+pG     + G+ C
     ratT272   322    CSNACTCKNGGTCVPENG---------NCVCAPG-----FRGPSC  352


laminin_EGF: domain 8 of 11, from 326 to 365: score -1.8, E = 0.67

                   *->CdCnphGslsddtCdsddelfgeetGqClkCkpnvtGrrCdr.CkpG
                      C+C + G     tC +       e G C+ C p++ G+ C r+C pG
     ratT272   326    CTCKNGG-----TCVP-------ENGNCV-CAPGFRGPSCQRpCPPG 359

                   yyglpsgdpgqgC<-*
                    y      + +  C
     ratT272   360 RY-----GKR-C     365


EGF: domain 9 of 11, from 365 to 394: score 18.3, E = 0.18

                   *->CapnnpCsngGtCvntpggssdnfggytCeCppGdyylsytGkrC<-
                      C p   C+n+  C+++ g         tC C  G     +tG++C
     ratT272   365    CVPC-KCNNHSSCHPSDG-------.---TCSCLAG-----WTGPDC   394
```

FIG.41C

```
laminin_EGF: domain 9 of 11, from 368 to 407: score 24.0, E = 0.0034

                   *->CdCnphGslsddtCdsddelfgeetGqClkCkpnvtGrrC.drCkpG
                      C Cn+h+      +C++        + G C+ +  + tG++C++ C pG
     ratT272   368    CKCNNHS-----SCHP-------SDGTCS-CLAGWTGPDCsESCPPG 401

                   yyglpsgdpgqgC<-*
                   ++gl          C
     ratT272   402 HWGL-------KC    407


EGF: domain 10 of 11, from 407 to 437: score 24.0, E = 0.035

                   *->CapnnpCsngGtCvntpggssdnfggytCeCppGdyylsytGkrC<-
                      C++++ C++g+tC++  g          +C+C pG      +tG++C
     ratT272   407    CSQPCQCHHGATCHPQDG---------SCVCIPG-----WTGPNC   437


laminin_EGF: domain 10 of 11, from 407 to 437: score 6.5, E = 0.12

                   *->CdCnphGslsddtCdsddelfgeetGqClkCkpnvtGrrCdrCkpGy
                      C+C++ +     tC++         G C+ C p+ tG++C +
     ratT272   411    CQCHHGA-----TCHP-------QDGSCV-CIPGWTGPNCSE----- 439

                   yglpsgdpgqgC<-*
                    g ps+++g++C
     ratT272   440 -GCPSRMFGVNC    450


EGF: domain 11 of 11, from 450 to 480: score 8.7, E = 3.7

                   *->CapnnpCsngGtCvntpggssdnfggytCeCppGdyylsytGkrC<-
                      C++++ C+ g  C++  g           C+CppG       +G +C
     ratT272   450    CSQLCQCDPGEMCHPETG---------ACVCPPG-----HSGAHC   480


laminin_EGF: domain 11 of 11, from 454 to 489: score -6.3, E = 1.7

                   *->CdCnphGslsddtCdsddelfgeetGqClkCkpnvtGrrCdrCkpGy
                      C+C+p G    + C++        etG+C+ C p+ +G +C
     ratT272   454    CQCDP-G----EMCHP-------ETGACV-CPPGHSGAHC------K 481

                   yglpsgdpgqgC<-*
                    g  + ++
     ratT272   482 VGSQE-SFT---    489
```

FIG.41D

SEQUENCE LISTING

<110> Millennium Pharmaceuticals, Inc.

<120> MEMBRANE-ASSOCIATED AND SECRETED PROTEINS AND USES THEREOF

<130> 7853-206-228

<150> 09/345,464
<151> 1999-06-30

<160> 148

<170> FastSEQ for Windows Version 3.0

<210> 1
<211> 3284
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (1222)...(1944)

<400> 1

```
gtcgacccac gcgtccgtta tgtaactata cattttccca gaaattttag tatatgatat     60
gattttgttt tctttcatcc cttttcccaa gcagtttatt atgaaaattt tcaaacatac    120
agcaatgttg agaaaatttt acagtaaatg cctataccca ttacctaaat tttaccatta    180
acattttacc ctgctggcat tattgtgctt atccatctac gtatccctct ctcccttcat    240
tggtgtattt ctaagtaaat tgtaggcctc agtacacttc cttctgaatt cttcagcatg    300
cacaacagta ttatattcca tttttaaaag agcaattctt gatagattta tatagttttg    360
taaaatgttc atatagagct acaaatttta tctttttgtt tcttattgta tgtctagggt    420
cctgaagggg atgctggcat tgttgggata tcaggtccta aaggtcctat tggacacaga    480
ggaaacactg gtccccttgg cagagaaggt ataataggcc caacaggtag aactggaccc    540
agaggtgaaa agggctttag aggtgaaact ggtcctcaag gaccaagagg tcaaccaggg    600
cctccaggtc cacctggagc accaggccca agaaagcaaa tggatatcaa tgctgctatt    660
caagccttga ttgaatcaaa tactgcccta cagatggagg taacatatct ggttttattt    720
atattggcac tgtctctcaa tataccaatt aaacagagaa aatttttgga ggccaaaatg    780
tgacattatc tcaaagattg tatttaaaac agattgaaaa tgtgaaacca ttctcaagaa    840
caaagtaagt gattttggta taattaaaca gaaatatatg cgtaggatgt tttgtaagga    900
aaacatttaa atcaaaaatt tagtactgtt atttgtaagg aatttggtac tatccaagaa    960
agtagttaaa tgaggttagc catgtttctt aaaatgagat atatatatta tcactactca   1020
tttatttaaa ctctaatgat tcaatgtgta atttaaaaaa cataatacag tagacatagc   1080
aattcttatg ttagcttgaa aactaaactt gcaaatgtga atttaacctc tttaaaagat   1140
taaggttatt aaagcataca catatgccta tgcttaaata taaactgttc tttacattct   1200
actcacaact tactacacat a atg gaa aca cat tct tct cct gcc ttg gcc      1251
                       Met Glu Thr His Ser Ser Pro Ala Leu Ala
                        1               5                  10

cat gtt ggt cct cag gat ttt ttt gtt tat ata att ctt atg atg act      1299
His Val Gly Pro Gln Asp Phe Phe Val Tyr Ile Ile Leu Met Met Thr
                15                  20                  25

tgg cag agc tac cag aat act gaa gtg act tta att gac cac agt gaa      1347
Trp Gln Ser Tyr Gln Asn Thr Glu Val Thr Leu Ile Asp His Ser Glu
            30                  35                  40
```

```
gag ata ttc aaa acc ctg aac tac ctt agc aat tta ttg cac agc atc     1395
Glu Ile Phe Lys Thr Leu Asn Tyr Leu Ser Asn Leu Leu His Ser Ile
         45                  50                  55

aag aat cct ctt ggc aca cga gat aac cca gca cga atc tgc aaa gat     1443
Lys Asn Pro Leu Gly Thr Arg Asp Asn Pro Ala Arg Ile Cys Lys Asp
     60                  65                  70

tta ctt aac tgt gaa caa aaa gta tca gat gga aaa tac tgg att gac     1491
Leu Leu Asn Cys Glu Gln Lys Val Ser Asp Gly Lys Tyr Trp Ile Asp
 75                  80                  85                  90

cca aat ctt ggc tgt cct tca gat gcc att gag gtt ttc tgc aat ttc     1539
Pro Asn Leu Gly Cys Pro Ser Asp Ala Ile Glu Val Phe Cys Asn Phe
                 95                 100                 105

agt gct ggt ggc cag aca tgc tta cct cct gtt tct gta aca aag ttg     1587
Ser Ala Gly Gly Gln Thr Cys Leu Pro Pro Val Ser Val Thr Lys Leu
            110                 115                 120

gag ttt gga gtt ggg aaa gtc cag atg aac ttc ctt cat tta ctg agt     1635
Glu Phe Gly Val Gly Lys Val Gln Met Asn Phe Leu His Leu Leu Ser
        125                 130                 135

tcg gaa gcc acc cat atc atc acc att cac tgt cta aac acc cca agg     1683
Ser Glu Ala Thr His Ile Ile Thr Ile His Cys Leu Asn Thr Pro Arg
    140                 145                 150

tgg aca agc aca caa aca agt ggc cca gga ttg cct att ggt ttc aag     1731
Trp Thr Ser Thr Gln Thr Ser Gly Pro Gly Leu Pro Ile Gly Phe Lys
155                 160                 165                 170

gga tgg aat ggc cag att ttt aaa gta aac act cta ctt gaa cct aaa     1779
Gly Trp Asn Gly Gln Ile Phe Lys Val Asn Thr Leu Leu Glu Pro Lys
                175                 180                 185

gtg ctt tca gat gac tgc aag att caa gat ggc agc tgg cat aag gca     1827
Val Leu Ser Asp Asp Cys Lys Ile Gln Asp Gly Ser Trp His Lys Ala
            190                 195                 200

aca ttt ctt ttt cac acc cag gaa cct aat caa ctt cca gtg att gaa     1875
Thr Phe Leu Phe His Thr Gln Glu Pro Asn Gln Leu Pro Val Ile Glu
        205                 210                 215

gta caa aaa ctt cct cat ctc aaa act gaa cga aag tat tac att gac     1923
Val Gln Lys Leu Pro His Leu Lys Thr Glu Arg Lys Tyr Tyr Ile Asp
    220                 225                 230

agc agt tct gta tgc ttt ctg taaagtctct gaattagttc cgaattcagg        1974
Ser Ser Ser Val Cys Phe Leu
235                 240

ctgttggcca ggtaattgct gcagagggag aaataagaca gacagataca gtcattatga   2034
aatgcatgta ataaagcatt ggctaaatct taaagaatct caggaagaac agacttcctc   2094
ctaagaagga gaaaaggcat ttttaaagga ctatgattga taaagtattt aattctttta   2154
aaaattatat tcatctcagc tttcttagag aattccctag aactaaaaat ttataaatat   2214
ggaattcttc agggtatctt atatttttga ctgagtgcgt agtacccatt agacagctgg   2274
agatgcagag cactatggag caatactggc taatgcttcc agatgtgcac tgcttctgtc   2334
```

taaaaattac aagccacagt ctaatatgtc ttattttcca aaacactaag ctgtattcag 2394
gtccccgatg ggcatataca tcttagccgg tgatacacta cctcttacgt gttgcctctt 2454
tgtgttgctt ggtgctcttt cgaaaacaag gtgcttatgg ctttcataga ctatttcctt 2514
tttcatcttt gtcattcttt aaaagtgtat gtactggtta catcaagata tgttttggtt 2574
gttagtactt attttaattt gtttggtcac acacttaata acacatgaaa ctatttatgt 2634
gaagtccttg ttttatttta aaattctctt tgtgtatttg gaatcaaagc cagcacattg 2694
taacctgtgc ttgtacgcaa aagaattaga tttctttgtt tttgttttat tttttaaatt 2754
gttgtaaaaa ttattatagg ccagctacat ctagtagtag gtttggggta cagattgggg 2814
gttgtgccat actgttttta aagttcatga tcatctggaa tgatacttag tgtatatata 2874
ttttgtaaag ttttaattca gcaaattttt tgaaattgct gctgttttaa attataaaac 2934
ctttatattt ctgctttgta gaaattatat gttttgtagt attcattgat tttctttcac 2994
tgtacttaaa tttagtgtta gtactttaaa atttttaatt taccagtctt taaagcaaca 3054
tccagaaaaa aaaaagtctt ttcccattta aaataggctc agccagttca atgtcgcctt 3114
gttatcagag aaatattagt tcaatactga aagaaaaata ttatacctct tggtatctag 3174
aaaagcttgt tcatccatta taaatatatc tttagccaca gcaaaccaca cttaacctat 3234
ctataataaa aatgtgcttt aaataaaaaa aaaaaaaaaa agggcggccg 3284

<210> 2
<211> 241
<212> PRT
<213> Homo sapiens

<400> 2

```
Met Glu Thr His Ser Ser Pro Ala Leu Ala His Val Gly Pro Gln Asp
 1               5                  10                  15
Phe Phe Val Tyr Ile Ile Leu Met Met Thr Trp Gln Ser Tyr Gln Asn
            20                  25                  30
Thr Glu Val Thr Leu Ile Asp His Ser Glu Glu Ile Phe Lys Thr Leu
        35                  40                  45
Asn Tyr Leu Ser Asn Leu Leu His Ser Ile Lys Asn Pro Leu Gly Thr
    50                  55                  60
Arg Asp Asn Pro Ala Arg Ile Cys Lys Asp Leu Leu Asn Cys Glu Gln
65                  70                  75                  80
Lys Val Ser Asp Gly Lys Tyr Trp Ile Asp Pro Asn Leu Gly Cys Pro
                85                  90                  95
Ser Asp Ala Ile Glu Val Phe Cys Asn Phe Ser Ala Gly Gly Gln Thr
            100                 105                 110
Cys Leu Pro Pro Val Ser Val Thr Lys Leu Glu Phe Gly Val Gly Lys
        115                 120                 125
Val Gln Met Asn Phe Leu His Leu Leu Ser Ser Glu Ala Thr His Ile
    130                 135                 140
Ile Thr Ile His Cys Leu Asn Thr Pro Arg Trp Thr Ser Thr Gln Thr
145                 150                 155                 160
Ser Gly Pro Gly Leu Pro Ile Gly Phe Lys Gly Trp Asn Gly Gln Ile
                165                 170                 175
Phe Lys Val Asn Thr Leu Leu Glu Pro Lys Val Leu Ser Asp Asp Cys
            180                 185                 190
Lys Ile Gln Asp Gly Ser Trp His Lys Ala Thr Phe Leu Phe His Thr
        195                 200                 205
Gln Glu Pro Asn Gln Leu Pro Val Ile Glu Val Gln Lys Leu Pro His
    210                 215                 220
Leu Lys Thr Glu Arg Lys Tyr Tyr Ile Asp Ser Ser Ser Val Cys Phe
225                 230                 235                 240
Leu
```

<210> 3
<211> 723

<212> DNA
<213> Homo sapiens

<400> 3
atggaaacac attcttctcc tgccttggcc catgttggtc ctcaggattt ttttgtttat 60
ataattctta tgatgacttg gcagagctac cagaatactg aagtgacttt aattgaccac 120
agtgaagaga tattcaaaac cctgaactac cttagcaatt tattgcacag catcaagaat 180
cctcttggca cacgagataa cccagcacga atctgcaaag atttacttaa ctgtgaacaa 240
aaagtatcag atggaaaata ctggattgac ccaaatcttg gctgtcctcc agatgccatt 300
gaggttttct gcaatttcag tgctggtggc cagacatgct tacctcctgt ttctgtaaca 360
aagttggagt ttggagttgg gaaagtccag atgaacttcc ttcatttact gagttcggaa 420
gccacccata tcatcaccat tcactgtcta aacaccccaa ggtggacaag cacacaaaca 480
agtggcccag gattgcctat tggtttcaag ggatggaatg gccagatttt taaagtaaac 540
actctacttg aacctaaagt gctttcagat gactgcaaga ttcaagatgg cagctggcat 600
aaggcaacat ttctttttca cacccaggaa cctaatcaac ttccagtgat tgaagtacaa 660
aaacttcctc atctcaaaac tgaacgaaag tattacattg acagcagttc tgtatgcttt 720
ctg 723

<210> 4
<211> 3169
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (57)...(1568)

<400> 4
gtcgacccac gcgtccgcgc ccgctgagcc ccccgccgag gtccggacag gccgag atg 59
Met
1

acg ccg agc ccc ctg ttg ctg ctc ctg ctg ccg ccg ctg ctg ctg ggg 107
Thr Pro Ser Pro Leu Leu Leu Leu Leu Leu Pro Pro Leu Leu Leu Gly
5 10 15

gcc ttc ccg ccg gcc gcc gcc gcc cga ggc ccc cca aag atg gcg gac 155
Ala Phe Pro Pro Ala Ala Ala Ala Arg Gly Pro Pro Lys Met Ala Asp
20 25 30

aag gtg gtc cca cgg cag gtg gcc cgg ctg ggc cgc act gtg cgg ctg 203
Lys Val Val Pro Arg Gln Val Ala Arg Leu Gly Arg Thr Val Arg Leu
35 40 45

cag tgc cca gtg gag ggg gac ccg ccg ccg ctg acc atg tgg acc aag 251
Gln Cys Pro Val Glu Gly Asp Pro Pro Pro Leu Thr Met Trp Thr Lys
50 55 60 65

gat ggc cgc acc atc cac agc ggc tgg agc cgc ttc cgc gtg ctg ccg 299
Asp Gly Arg Thr Ile His Ser Gly Trp Ser Arg Phe Arg Val Leu Pro
70 75 80

cag ggg ctg aag gtg aag cag gtg gag cgg gag gat gcc ggc gtg tac 347
Gln Gly Leu Lys Val Lys Gln Val Glu Arg Glu Asp Ala Gly Val Tyr
85 90 95

gtg tgc aag gcc acc aac ggc ttc ggc agc ctg agc gtc aac tac acc 395
Val Cys Lys Ala Thr Asn Gly Phe Gly Ser Leu Ser Val Asn Tyr Thr

```
        100                 105                 110

ctc gtc gtg ctg gat gac att agc cca ggg aag gag agc ctg ggg ccc      443
Leu Val Val Leu Asp Asp Ile Ser Pro Gly Lys Glu Ser Leu Gly Pro
    115                 120                 125

gac agc tcc tct ggg ggt caa gag gac ccc gcc agc cag cag tgg gca      491
Asp Ser Ser Ser Gly Gly Gln Glu Asp Pro Ala Ser Gln Gln Trp Ala
130                 135                 140                 145

cga ccg cgc ttc aca cag ccc tcc aag atg agg cgc cgg gtg atc gca      539
Arg Pro Arg Phe Thr Gln Pro Ser Lys Met Arg Arg Arg Val Ile Ala
                150                 155                 160

cgg ccc gtg ggt agc tcc gtg cgg ctc aag tgc gtg gcc agc ggg cac      587
Arg Pro Val Gly Ser Ser Val Arg Leu Lys Cys Val Ala Ser Gly His
            165                 170                 175

cct cgg ccc gac atc acg tgg atg aag gac gac cag gcc ttg acg cgc      635
Pro Arg Pro Asp Ile Thr Trp Met Lys Asp Asp Gln Ala Leu Thr Arg
        180                 185                 190

cca gag gcc gct gag ccc agg aag aag aag tgg aca ctg agc ctg aag      683
Pro Glu Ala Ala Glu Pro Arg Lys Lys Lys Trp Thr Leu Ser Leu Lys
    195                 200                 205

aac ctg cgg ccg gag gac agc ggc aaa tac acc tgc cgc gtg tcg aac      731
Asn Leu Arg Pro Glu Asp Ser Gly Lys Tyr Thr Cys Arg Val Ser Asn
210                 215                 220                 225

cgc gcg ggc gcc atc aac gcc acc tac aag gtg gat gtg atc cag cgg      779
Arg Ala Gly Ala Ile Asn Ala Thr Tyr Lys Val Asp Val Ile Gln Arg
                230                 235                 240

acc cgt tcc aag ccc gtg ctc aca ggc acg cac ccc gtg aac acg acg      827
Thr Arg Ser Lys Pro Val Leu Thr Gly Thr His Pro Val Asn Thr Thr
            245                 250                 255

gtg gac ttc ggg ggg acc acg tcc ttc cag tgc aag gtg cgc agc gac      875
Val Asp Phe Gly Gly Thr Thr Ser Phe Gln Cys Lys Val Arg Ser Asp
        260                 265                 270

gtg aag ccg gtg atc cag tgg ctg aag cgc gtg gag tac ggc gcc gag      923
Val Lys Pro Val Ile Gln Trp Leu Lys Arg Val Glu Tyr Gly Ala Glu
    275                 280                 285

ggc cgc cac aac tcc acc atc gat gtg ggc ggc cag aag ttt gtg gtg      971
Gly Arg His Asn Ser Thr Ile Asp Val Gly Gly Gln Lys Phe Val Val
290                 295                 300                 305

ctg ccc acg ggt gac gtg tgg tcg cgg ccc gac ggc tcc tac ctc aat     1019
Leu Pro Thr Gly Asp Val Trp Ser Arg Pro Asp Gly Ser Tyr Leu Asn
                310                 315                 320

aag ctg ctc atc acc cgt gcc cgc cag gac gat gcg ggc atg tac atc     1067
Lys Leu Leu Ile Thr Arg Ala Arg Gln Asp Asp Ala Gly Met Tyr Ile
            325                 330                 335
```

```
tgc ctt ggc gcc aac acc atg ggc tac agc ttc cgc agc gcc ttc ctc     1115
Cys Leu Gly Ala Asn Thr Met Gly Tyr Ser Phe Arg Ser Ala Phe Leu
        340                 345                 350

acc gtg ctg cca gac cca aaa ccg cca ggg cca cct gtg gcc tcc tcg     1163
Thr Val Leu Pro Asp Pro Lys Pro Pro Gly Pro Pro Val Ala Ser Ser
    355                 360                 365

tcc tcg gcc act agc ctg ccg tgg ccc gtg gtc atc ggc atc cca gcc     1211
Ser Ser Ala Thr Ser Leu Pro Trp Pro Val Val Ile Gly Ile Pro Ala
370                 375                 380                 385

ggc gct gtc ttc atc ctg ggc acc ctg ctc ctg tgg ctt tgc cag gcc     1259
Gly Ala Val Phe Ile Leu Gly Thr Leu Leu Leu Trp Leu Cys Gln Ala
                390                 395                 400

cag aag aag ccg tgc acc ccc gcg cct gcc cct ccc ctg cct ggg cac     1307
Gln Lys Lys Pro Cys Thr Pro Ala Pro Ala Pro Pro Leu Pro Gly His
            405                 410                 415

cgc ccg ccg ggg acg gcc cgc gac cgc agc gga gac aag gac ctt ccc     1355
Arg Pro Pro Gly Thr Ala Arg Asp Arg Ser Gly Asp Lys Asp Leu Pro
        420                 425                 430

tcg ttg gcc gcc ctc agc gct ggc cct ggt gtg ggg ctg tgt gag gag     1403
Ser Leu Ala Ala Leu Ser Ala Gly Pro Gly Val Gly Leu Cys Glu Glu
    435                 440                 445

cat ggg tct ccg gca gcc ccc cag cac tta ctg ggc cca ggc cca gtt     1451
His Gly Ser Pro Ala Ala Pro Gln His Leu Leu Gly Pro Gly Pro Val
450                 455                 460                 465

gct ggc cct aag ttg tac ccc aaa ctc tac aca gac atc cac aca cac     1499
Ala Gly Pro Lys Leu Tyr Pro Lys Leu Tyr Thr Asp Ile His Thr His
                470                 475                 480

aca cac aca cac tct cac aca cac tca cac gtg gag ggc aag gtc cac     1547
Thr His Thr His Ser His Thr His Ser His Val Glu Gly Lys Val His
            485                 490                 495

cag cac atc cac tat cag tgc tagacggcac cgtatctgca gtgggcacgg        1598
Gln His Ile His Tyr Gln Cys
        500

gggggccggc cagacaggca gactgggagg atggaggacg gagctgcaga cgaaggcagg   1658
ggacccatgg cgaggaggaa tggccagcac cccaggcagt ctgtgtgtga ggcatagccc   1718
ctggacacac acacacagac acacacactg cctggatgca tgtatgcaca cacatgcgcg   1778
cacacgtgct ccctgaaggc acacgtacgc acacacgcac atgcacagat atgccgcctg   1838
ggcacacaga taagctgccc aaatgcacgc acacgcacag agacatgcca gaacatacaa   1898
ggacatgctg cctgaacata cacacgcaca cccatgcgca gatgtgctgc ctggacacac   1958
acacacacac ggatatgctg tctggacgca cacacgtgca gatatggtat ccggacacac   2018
acgtgcacag atatgctgcc tggacacaca gataatgctg ccttgacaca cacatgcacg   2078
gatattgcct ggacacacac acacacacgt gtgcacagat atgctgtctg gacacgcaca   2138
cacatgcaga tatgctgcct ggacacacac ttccagacac acgtgcacag gcgcagatat   2198
gctgcctgga cacacgcaga tatgctgtct agtcacacac acacgcagac atgctgtccg   2258
gacacacaca cgcatgcaca gatatgctgt ccggacacac acacgcacgc agatatgctg   2318
cctggacaca cacacagata atgctgcctc aacactcaca cacgtgcaga tattgcctgg   2378
acacacacat gtgcacagat atgctgtctg gacatgcaca cacgtgcaga tatgctgtcc   2438
```

```
ggatacacac gcacgcacac atgcagatat gctgcctggg cacacacttc cggacacaca   2498
tgcacacaca ggtgcagata tgctgcctgg acacacgcag actgacgtgc ttttgggagg   2558
gtgtgccgtg aagcctgcag tacgtgtgcc gtgaggctca tagttgatga gggactttcc   2618
ctgctccacc gtcactcccc caactctgcc cgcctctgtc cccgcctcag tccccgcctc   2678
catccccgcc tctgtcccct ggccttggcg gctatttttg ccacctgcct tgggtgccca   2738
ggagtccccc actgctgtgg gctggggttg ggggcacagc agccccaagc ctgagaggct   2798
ggagcccatg gctagtggct catccccact gcattctccc cctgacacag agaaggggcc   2858
ttggtattta tatttaagaa atgaagataa tattaataat gatggaagga agactgggtt   2918
gcagggactg tggtctctcc tggggcccgg gacccgcctg gtctttcagc catgctgatg   2978
accacacccc gtccaggcca gacaccaccc cccaccccac tgtcgtggtg gccccagatc   3038
tctgtaattt tatgtagagt ttgagctgaa gccccgtata tttaatttat tttgttaaac   3098
atgaaagtgc aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa   3158
agggcggccg c                                                       3169
```

<210> 5
<211> 504
<212> PRT
<213> Homo sapiens

<400> 5

```
Met Thr Pro Ser Pro Leu Leu Leu Leu Leu Leu Pro Pro Leu Leu Leu
 1               5                  10                  15
Gly Ala Phe Pro Pro Ala Ala Ala Ala Arg Gly Pro Pro Lys Met Ala
            20                  25                  30
Asp Lys Val Val Pro Arg Gln Val Ala Arg Leu Gly Arg Thr Val Arg
        35                  40                  45
Leu Gln Cys Pro Val Glu Gly Asp Pro Pro Pro Leu Thr Met Trp Thr
    50                  55                  60
Lys Asp Gly Arg Thr Ile His Ser Gly Trp Ser Arg Phe Arg Val Leu
65                  70                  75                  80
Pro Gln Gly Leu Lys Val Lys Gln Val Glu Arg Glu Asp Ala Gly Val
                85                  90                  95
Tyr Val Cys Lys Ala Thr Asn Gly Phe Gly Ser Leu Ser Val Asn Tyr
            100                 105                 110
Thr Leu Val Val Leu Asp Asp Ile Ser Pro Gly Lys Glu Ser Leu Gly
        115                 120                 125
Pro Asp Ser Ser Ser Gly Gly Gln Glu Asp Pro Ala Ser Gln Gln Trp
    130                 135                 140
Ala Arg Pro Arg Phe Thr Gln Pro Ser Lys Met Arg Arg Arg Val Ile
145                 150                 155                 160
Ala Arg Pro Val Gly Ser Ser Val Arg Leu Lys Cys Val Ala Ser Gly
                165                 170                 175
His Pro Arg Pro Asp Ile Thr Trp Met Lys Asp Asp Gln Ala Leu Thr
            180                 185                 190
Arg Pro Glu Ala Ala Glu Pro Arg Lys Lys Lys Trp Thr Leu Ser Leu
        195                 200                 205
Lys Asn Leu Arg Pro Glu Asp Ser Gly Lys Tyr Thr Cys Arg Val Ser
    210                 215                 220
Asn Arg Ala Gly Ala Ile Asn Ala Thr Tyr Lys Val Asp Val Ile Gln
225                 230                 235                 240
Arg Thr Arg Ser Lys Pro Val Leu Thr Gly Thr His Pro Val Asn Thr
                245                 250                 255
Thr Val Asp Phe Gly Gly Thr Thr Ser Phe Gln Cys Lys Val Arg Ser
            260                 265                 270
Asp Val Lys Pro Val Ile Gln Trp Leu Lys Arg Val Glu Tyr Gly Ala
        275                 280                 285
Glu Gly Arg His Asn Ser Thr Ile Asp Val Gly Gly Gln Lys Phe Val
    290                 295                 300
```

```
Val Leu Pro Thr Gly Asp Val Trp Ser Arg Pro Asp Gly Ser Tyr Leu
305                 310                 315                 320
Asn Lys Leu Leu Ile Thr Arg Ala Arg Gln Asp Asp Ala Gly Met Tyr
                325                 330                 335
Ile Cys Leu Gly Ala Asn Thr Met Gly Tyr Ser Phe Arg Ser Ala Phe
            340                 345                 350
Leu Thr Val Leu Pro Asp Pro Lys Pro Pro Gly Pro Pro Val Ala Ser
        355                 360                 365
Ser Ser Ser Ala Thr Ser Leu Pro Trp Pro Val Val Ile Gly Ile Pro
    370                 375                 380
Ala Gly Ala Val Phe Ile Leu Gly Thr Leu Leu Leu Trp Leu Cys Gln
385                 390                 395                 400
Ala Gln Lys Lys Pro Cys Thr Pro Ala Pro Ala Pro Pro Leu Pro Gly
                405                 410                 415
His Arg Pro Pro Gly Thr Ala Arg Asp Arg Ser Gly Asp Lys Asp Leu
            420                 425                 430
Pro Ser Leu Ala Ala Leu Ser Ala Gly Pro Gly Val Gly Leu Cys Glu
        435                 440                 445
Glu His Gly Ser Pro Ala Ala Pro Gln His Leu Leu Gly Pro Gly Pro
    450                 455                 460
Val Ala Gly Pro Lys Leu Tyr Pro Lys Leu Tyr Thr Asp Ile His Thr
465                 470                 475                 480
His Thr His Thr His Ser His Thr His Ser His Val Glu Gly Lys Val
                485                 490                 495
His Gln His Ile His Tyr Gln Cys
            500
```

<210> 6
<211> 1512
<212> DNA
<213> Homo sapiens

<400> 6

```
atgacgccga gccccctgtt gctgctcctg ctgccgccgc tgctgctggg ggccttcccg      60
ccggccgccg ccgcccgagg ccccccaaag atggcggaca aggtggtccc acggcaggtg     120
gcccggctgg gccgcactgt gcggctgcag tgcccagtgg agggggaccc gccgccgctg     180
accatgtgga ccaaggatgg ccgcaccatc cacagcggct ggagccgctt ccgcgtgctg     240
ccgcaggggc tgaaggtgaa gcaggtggag cgggaggatg ccggcgtgta cgtgtgcaag     300
gccaccaacg gcttcggcag cctgagcgtc aactacaccc tcgtcgtgct ggatgacatt     360
agcccaggga aggagagcct ggggcccgac agctcctctg ggggtcaaga ggaccccgcc     420
agccagcagt gggcacgacc gcgcttcaca cagccctcca agatgaggcg ccgggtgatc     480
gcacggcccg tgggtagctc cgtgcggctc aagtgcgtgg ccagcgggca ccctcggccc     540
gacatcacgt ggatgaagga cgaccaggcc ttgacgcgcc cagaggccgc tgagcccagg     600
aagaagaagt ggacactgag cctgaagaac ctgcggccgg aggacagcgg caaatacacc     660
tgccgcgtgt cgaaccgcgc gggcgccatc aacgccacct acaaggtgga tgtgatccag     720
cggacccgtt ccaagcccgt gctcacaggc acgcaccccg tgaacacgac ggtggacttc     780
gggggacca cgtccttcca gtgcaaggtg cgcagcgacg tgaagccggt gatccagtgg     840
ctgaagcgcg tggagtacgg cgccgagggc cgccacaact ccaccatcga tgtgggcggc     900
cagaagtttg tggtgctgcc cacgggtgac gtgtggtcgc ggcccgacgg ctcctacctc     960
aataagctgc tcatcacccg tgcccgccag gacgatgcgg gcatgtacat ctgccttggc    1020
gccaacacca tgggctacag cttccgcagc gccttcctca ccgtgctgcc agacccaaaa    1080
ccgccagggc cacctgtggc ctcctcgtcc tcggccacta gcctgccgtg gcccgtggtc    1140
atcggcatcc cagccggcgc tgtcttcatc ctgggcaccc tgctcctgtg gctttgccag    1200
gcccagaaga agccgtgcac ccccgcgcct gcccctcccc tgcctgggca ccgcccgccg    1260
gggacggccc gcgaccgcag cggagacaag gaccttccct cgttggccgc cctcagcgct    1320
ggccctggtg tggggctgtg tgaggagcat gggtctccgg cagccccca gcacttactg    1380
ggcccaggcc cagttgctgg ccctaagttg taccccaaac tctacacaga catccacaca    1440
cacacacaca cacactctca cacacactca cacgtggagg gcaaggtcca ccagcacatc    1500
```

```
cactatcagt gc                                                          1512

      <210> 7
      <211> 1074
      <212> DNA
      <213> Mus musculus

      <220>
      <221> CDS
      <222> (3)...(626)

      <221> modified_base
      <222> all "n" positions
      <223> n=a, c, g, or t

      <400> 7
ca cgc gtc cgg ccc acg ggt gat gtg tgg tca cgg cct gat ggc tcc          47
   Arg Val Arg Pro Thr Gly Asp Val Trp Ser Arg Pro Asp Gly Ser
    1               5                  10                  15

tac ctc aac aag ctg ctc atc tct cgg gcc cgc cag gat gat gct ggc          95
Tyr Leu Asn Lys Leu Leu Ile Ser Arg Ala Arg Gln Asp Asp Ala Gly
                20                  25                  30

atg tac atc tgc cta ggt gca aat acc atg ggc tac agt ttc cgt agc         143
Met Tyr Ile Cys Leu Gly Ala Asn Thr Met Gly Tyr Ser Phe Arg Ser
            35                  40                  45

gcc ttc ctc act gta tta cca gac ccc aaa cct cca ggg cct cct atg         191
Ala Phe Leu Thr Val Leu Pro Asp Pro Lys Pro Pro Gly Pro Pro Met
        50                  55                  60

gct tct tca tcg tca tcc aca agc ctg cca tgg cct gtg gtg atc ggc         239
Ala Ser Ser Ser Ser Ser Thr Ser Leu Pro Trp Pro Val Val Ile Gly
    65                  70                  75

atc cca gct ggt gct gtc ttc atc cta ggc act gtg ctg ctc tgg ctt         287
Ile Pro Ala Gly Ala Val Phe Ile Leu Gly Thr Val Leu Leu Trp Leu
80                  85                  90                  95

tgc cag acc aag aag aag cca tgt gcc cca gca tct aca ctt cct gtg         335
Cys Gln Thr Lys Lys Lys Pro Cys Ala Pro Ala Ser Thr Leu Pro Val
                100                 105                 110

cct ggg cat cgt ccc cca ggg aca tcc cga gaa cgc agt ggt gac aag         383
Pro Gly His Arg Pro Pro Gly Thr Ser Arg Glu Arg Ser Gly Asp Lys
            115                 120                 125

gac ctg ccc tca ttg gct gtg ggc ata tgt gag gag cat gga tcc gcc         431
Asp Leu Pro Ser Leu Ala Val Gly Ile Cys Glu Glu His Gly Ser Ala
        130                 135                 140

atg gcc ccc cag cac atc ctg gcc tct ggc tca act gct ggc ccc aag         479
Met Ala Pro Gln His Ile Leu Ala Ser Gly Ser Thr Ala Gly Pro Lys
    145                 150                 155

ctg tac ccc aag cta tac aca gat gtg cac aca cac aca cat aca cac         527
Leu Tyr Pro Lys Leu Tyr Thr Asp Val His Thr His Thr His Thr His
```

```
160                 165                 170                 175

acc tgc act cac acg ctc tca tgt tgg agg gca agg ttc atc aac acc      575
Thr Cys Thr His Thr Leu Ser Cys Trp Arg Ala Arg Phe Ile Asn Thr
                180                 185                 190

agc atg tcc act atc agt gct aaa tac agc gaa tct cca agc act gtg      623
Ser Met Ser Thr Ile Ser Ala Lys Tyr Ser Glu Ser Pro Ser Thr Val
            195                 200                 205

tcc tgaggtaggc atttgggggc caaggcaaca ggttgggaga attgagaaca          676
Ser


atggaggaag agtatcttag ggtgccttat ggtggacact cacaaacttg gccatataga    736
tgtatgtact accagatgaa cagccagcca gattcacaca cgcacatgtt taaacgtgta    796
aacgtgtgca caactgcaca cacaacctga gaaaccttca ggaggatttg tggtgtgact    856
ttgcagtgac atgtagcgat ggctagttga aggaatctcc ctcatgtctt agtggtcatg    916
gccacttccc caccctgcc catctgtgtt cctgcctggc cttggtggtg cttccgtgtg    976
ccctgggttt tccaggaacc ctatcaacct gactggggtg agcagtgcag ccatgcntgg   1036
aggtttgagc caccctcccc ttgctagaga gaagggcn                           1074

      <210> 8
      <211> 208
      <212> PRT
      <213> Mus musculus

      <400> 8
Arg Val Arg Pro Thr Gly Asp Val Trp Ser Arg Pro Asp Gly Ser Tyr
 1               5                  10                  15
Leu Asn Lys Leu Leu Ile Ser Arg Ala Arg Gln Asp Asp Ala Gly Met
            20                  25                  30
Tyr Ile Cys Leu Gly Ala Asn Thr Met Gly Tyr Ser Phe Arg Ser Ala
        35                  40                  45
Phe Leu Thr Val Leu Pro Asp Pro Lys Pro Pro Gly Pro Pro Met Ala
    50                  55                  60
Ser Ser Ser Ser Ser Thr Ser Leu Pro Trp Pro Val Val Ile Gly Ile
65                  70                  75                  80
Pro Ala Gly Ala Val Phe Ile Leu Gly Thr Val Leu Leu Trp Leu Cys
                85                  90                  95
Gln Thr Lys Lys Lys Pro Cys Ala Pro Ala Ser Thr Leu Pro Val Pro
            100                 105                 110
Gly His Arg Pro Pro Gly Thr Ser Arg Glu Arg Ser Gly Asp Lys Asp
        115                 120                 125
Leu Pro Ser Leu Ala Val Gly Ile Cys Glu Glu His Gly Ser Ala Met
    130                 135                 140
Ala Pro Gln His Ile Leu Ala Ser Gly Ser Thr Ala Gly Pro Lys Leu
145                 150                 155                 160
Tyr Pro Lys Leu Tyr Thr Asp Val His Thr His Thr His Thr His Thr
                165                 170                 175
Cys Thr His Thr Leu Ser Cys Trp Arg Ala Arg Phe Ile Asn Thr Ser
            180                 185                 190
Met Ser Thr Ile Ser Ala Lys Tyr Ser Glu Ser Pro Ser Thr Val Ser
        195                 200                 205

      <210> 9
      <211> 624
      <212> DNA
```

<213> Mus musculus

<400> 9
cgcgtccggc ccacgggtga tgtgtggtca cggcctgatg gctcctacct caacaagctg 60
ctcatctctc gggcccgcca ggatgatgct ggcatgtaca tctgcctagg tgcaaatacc 120
atgggctaca gtttccgtag cgccttcctc actgtattac cagaccccaa acctccaggg 180
cctcctatgg cttcttcatc gtcatccaca agcctgccat ggcctgtggt gatcggcatc 240
ccagctggtg ctgtcttcat cctaggcact gtgctgctct ggctttgcca gaccaagaag 300
aagccatgtg ccccagcatc tacacttcct gtgcctgggc atcgtccccc agggacatcc 360
cgagaacgca gtggtgacaa ggacctgccc tcattggctg tgggcatatg tgaggagcat 420
ggatccgcca tggcccccca gcacatcctg gcctctggct caactgctgg ccccaagctg 480
taccccaagc tatacacaga tgtgcacaca cacacacata cacacacctg cactcacacg 540
ctctcatgtt ggagggcaag gttcatcaac accagcatgt ccactatcag tgctaaatac 600
agcgaatctc caagcactgt gtcc 624

<210> 10
<211> 1423
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (31)...(444)

<400> 10
gtcgacccac gcgtccgccc acgcgtccgg atg cct gga ccc aga gtg tgg ggg 54
Met Pro Gly Pro Arg Val Trp Gly
1 5

aaa tat ctc tgg aga agc cct cac tcc aaa ggc tgt cca ggc gca atg 102
Lys Tyr Leu Trp Arg Ser Pro His Ser Lys Gly Cys Pro Gly Ala Met
10 15 20

tgg tgg ctg ctt ctc tgg gga gtc ctc cag gct tgc cca acc cgg ggc 150
Trp Trp Leu Leu Leu Trp Gly Val Leu Gln Ala Cys Pro Thr Arg Gly
25 30 35 40

tcc gtc ctc ttg gcc caa gag cta ccc cag cag ctg aca tcc ccc ggg 198
Ser Val Leu Leu Ala Gln Glu Leu Pro Gln Gln Leu Thr Ser Pro Gly
45 50 55

tac cca gag ccg tat ggc aaa ggc caa gag agc agc acg gac atc aag 246
Tyr Pro Glu Pro Tyr Gly Lys Gly Gln Glu Ser Ser Thr Asp Ile Lys
60 65 70

gct cca gag ggc ttt gct gtg agg ctc gtc ttc cag gac ttc gac ctg 294
Ala Pro Glu Gly Phe Ala Val Arg Leu Val Phe Gln Asp Phe Asp Leu
75 80 85

gag ccg tcc cag gac tgt gca ggg gac tct gtc aca gtg agc tgg gga 342
Glu Pro Ser Gln Asp Cys Ala Gly Asp Ser Val Thr Val Ser Trp Gly
90 95 100

tgg ggg ggg tcc cgc cag gac tgt ggc cag gga gat tcc cgg ggt tgt 390
Trp Gly Gly Ser Arg Gln Asp Cys Gly Gln Gly Asp Ser Arg Gly Cys
105 110 115 120

ggg aag tgg cgg tgc cct gaa tcc ccc atc tgg agg agg gat gaa ttt 438

```
Gly Lys Trp Arg Cys Pro Glu Ser Pro Ile Trp Arg Arg Asp Glu Phe
                125                 130                 135

tcc atg taggggcagt cgggcttggc ttaccgggga gcagtggtgg accccaggac      494
Ser Met

acagcctccc accagcgcct ccggggctgc catctgggcc ccacagagca aagagggcag    554
caagcaggcc ctgcgtttgg aaggcttatg aatggacaca caaatcttgc aaatctatgg    614
agccaggggc agggacgcac atattggttg ttaaaaatat gtcatcatgt atttgttgag    674
tgcctgctct atcaggtgag gaagctggac acaaataata acaaaagatt aagtcaccgt    734
tcacacttac cttggaagag ctattacaaa acttctaacg ccaaagcctt attcagaata    794
aggacatttt aaaaacagta cttgatggag tgatgcaagc ttgcagtccc agcagtatag    854
tcaggagact gaggctggag gatcagaggg ctggagccca gggttcaagg ccagcctaag    914
caacatagca agaccccatc tcaaaaataa gtaaataata aataaaaata aaaagagcac    974
attatctttt gatttaaatt ttatttatat caaaatgaca taaatttttg aactttattt   1034
tttaatttta aaatttttaa ttattatgga tacataatag ttgtaagact ttttgttttt   1094
taattaaagt tttctaaggc tgggcgcagt agctcatgtc tgtagtccca gcactttggg   1154
aggctgaggc gaaagaagca cttgagccca ggaatttgag accagcctgg gcaacatagc   1214
aagaccccat ctctacaaaa aaatttaaaa attagccaag tgtggtggca cgcacctgtg   1274
gtcccagcta caagggacgc tgaagtgaga ggatcacttg agcctggaag gtagaggctg   1334
cagtgagctc tgatcatgac accgtactcc agcctgggtg acagagtgag accctgtctc   1394
caaaaaaaaa aaaaaaaaag ggcggccgc                                     1423

      <210> 11
      <211> 138
      <212> PRT
      <213> Homo sapiens

      <400> 11
Met Pro Gly Pro Arg Val Trp Gly Lys Tyr Leu Trp Arg Ser Pro His
 1               5                  10                  15
Ser Lys Gly Cys Pro Gly Ala Met Trp Trp Leu Leu Leu Trp Gly Val
            20                  25                  30
Leu Gln Ala Cys Pro Thr Arg Gly Ser Val Leu Leu Ala Gln Glu Leu
        35                  40                  45
Pro Gln Gln Leu Thr Ser Pro Gly Tyr Pro Glu Pro Tyr Gly Lys Gly
    50                  55                  60
Gln Glu Ser Ser Thr Asp Ile Lys Ala Pro Glu Gly Phe Ala Val Arg
65                  70                  75                  80
Leu Val Phe Gln Asp Phe Asp Leu Glu Pro Ser Gln Asp Cys Ala Gly
                85                  90                  95
Asp Ser Val Thr Val Ser Trp Gly Trp Gly Gly Ser Arg Gln Asp Cys
            100                 105                 110
Gly Gln Gly Asp Ser Arg Gly Cys Gly Lys Trp Arg Cys Pro Glu Ser
        115                 120                 125
Pro Ile Trp Arg Arg Asp Glu Phe Ser Met
    130                 135

      <210> 12
      <211> 414
      <212> DNA
      <213> Homo sapiens

      <400> 12
atgcctggac ccagagtgtg ggggaaatat ctctggagaa gccctcactc caaaggctgt      60
ccaggcgcaa tgtggtggct gcttctctgg ggagtcctcc aggcttgccc aacccggggc     120
tccgtcctct tggcccaaga gctaccccag cagctgacat ccccgggtta cccagagccg     180
```

```
tatggcaaag gccaagagag cagcacggac atcaaggctc cagagggctt tgctgtgagg     240
ctcgtcttcc aggacttcga cctggagccg tcccaggact gtgcagggga ctctgtcaca     300
gtgagctggg gatgggggcg gtcccgccag gactgtggcc agggagattc ccggggttgt     360
gggaagtggc ggtgccctga atcccccatc tggaggaggg atgaattttc catg           414
```

<210> 13
<211> 5036
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (230)...(3379)

<400> 13

```
gtcgacccac gcgtccgctc gaagcgggga ccctcgcccc gtcctcggct gtccagtcct      60
cctcctcgca gacccccggcg gttcctaccc caggccgcag gggagacggt gccccaaggc     120
aggcttcata tcctgaacgc tgggatcccc caggacattc cctggccccc aggccccagg     180
tcccaggccc cagggctgag ctgtgggcag gccccacctg gcctctgca atg tca ccg     238
                                                      Met Ser Pro
                                                       1

cct ctg tgt ccc ctc ctt ctc ctg gct gtg ggc ctg cgg ctg gct gga      286
Pro Leu Cys Pro Leu Leu Leu Leu Ala Val Gly Leu Arg Leu Ala Gly
     5                  10                  15

act ctc aac ccc agt gat ccc aat acc tgc agc ttc tgg gaa agc ttc      334
Thr Leu Asn Pro Ser Asp Pro Asn Thr Cys Ser Phe Trp Glu Ser Phe
 20                  25                  30                  35

act acc acc acc aag gag tcc cac tcc cgc ccc ttc agc ctg ctc ccc      382
Thr Thr Thr Thr Lys Glu Ser His Ser Arg Pro Phe Ser Leu Leu Pro
                 40                  45                  50

tca gag ccc tgc gag cgg ccc tgg gag ggc ccc cat act tgc ccc agc      430
Ser Glu Pro Cys Glu Arg Pro Trp Glu Gly Pro His Thr Cys Pro Ser
            55                  60                  65

cca caa act cag agg aaa ctc ctg gct tct agg gat tca ttc tgc atg      478
Pro Gln Thr Gln Arg Lys Leu Leu Ala Ser Arg Asp Ser Phe Cys Met
        70                  75                  80

gtc tgt gtc ggg gct gga gtg cag tgg cga gat cgt agt gca ctg caa      526
Val Cys Val Gly Ala Gly Val Gln Trp Arg Asp Arg Ser Ala Leu Gln
     85                  90                  95

cct caa aca ggg aat gcg ctt tct atg cgc cct cag ccc aga gtg ttg      574
Pro Gln Thr Gly Asn Ala Leu Ser Met Arg Pro Gln Pro Arg Val Leu
100                 105                 110                 115

agt ggt gcc cct tcc ctg gcc tcc cct ggc cac act gtg gtg gtg aag      622
Ser Gly Ala Pro Ser Leu Ala Ser Pro Gly His Thr Val Val Val Lys
                120                 125                 130

acg gac cac cgc cag cgc ctg cag tgc tgc cat ggc ttc tat gag agc      670
Thr Asp His Arg Gln Arg Leu Gln Cys Cys His Gly Phe Tyr Glu Ser
            135                 140                 145
```

```
agg ggg ttc tgt gtc ccg ctc tgt gcc cag gag tgt gtc cat ggc cgt      718
Arg Gly Phe Cys Val Pro Leu Cys Ala Gln Glu Cys Val His Gly Arg
        150                 155                 160

tgt gtg gca ccc aat cag tgc caa tgt gtg cca ggc tgg cgg ggc gac      766
Cys Val Ala Pro Asn Gln Cys Gln Cys Val Pro Gly Trp Arg Gly Asp
    165                 170                 175

gac tgt tcc agt gcc ccg aac tgc ctt cag ccc tgt acc cct ggc tac      814
Asp Cys Ser Ser Ala Pro Asn Cys Leu Gln Pro Cys Thr Pro Gly Tyr
180                 185                 190                 195

tat ggc cct gcc tgc cag ttc cgc tgc cag tgc cat ggg gca ccc tgc      862
Tyr Gly Pro Ala Cys Gln Phe Arg Cys Gln Cys His Gly Ala Pro Cys
                200                 205                 210

gat ccc cag act gga gcc tgc ttc tgc ccc gca gag aga act ggg ccc      910
Asp Pro Gln Thr Gly Ala Cys Phe Cys Pro Ala Glu Arg Thr Gly Pro
            215                 220                 225

agc tgt gac gtg tcc tgt tcc cag ggc act tct ggc ttc ttc tgc ccc      958
Ser Cys Asp Val Ser Cys Ser Gln Gly Thr Ser Gly Phe Phe Cys Pro
        230                 235                 240

agc acc cat cct tgc caa aat gga ggt gtc ttc caa acc cca cag ggc     1006
Ser Thr His Pro Cys Gln Asn Gly Gly Val Phe Gln Thr Pro Gln Gly
    245                 250                 255

tcc tgc agc tgc ccc cct ggc tgg atg ggc acc atc tgc tcc ctg ccc     1054
Ser Cys Ser Cys Pro Pro Gly Trp Met Gly Thr Ile Cys Ser Leu Pro
260                 265                 270                 275

tgc cca gag ggc ttt cac gga ccc aac tgc tcc cag gaa tgt cgc tgc     1102
Cys Pro Glu Gly Phe His Gly Pro Asn Cys Ser Gln Glu Cys Arg Cys
                280                 285                 290

cac aac ggc ggc ctc tgt gac cga ttc act ggg cag tgc cgc tgc gct     1150
His Asn Gly Gly Leu Cys Asp Arg Phe Thr Gly Gln Cys Arg Cys Ala
            295                 300                 305

ccg ggt tac act ggg gat cgg tgc cgg gag gag tgc ccg gtg ggc cgc     1198
Pro Gly Tyr Thr Gly Asp Arg Cys Arg Glu Glu Cys Pro Val Gly Arg
        310                 315                 320

ttt ggg cag gac tgt gct gag acg tgc gac tgc gcc ccg gac gcc cgt     1246
Phe Gly Gln Asp Cys Ala Glu Thr Cys Asp Cys Ala Pro Asp Ala Arg
    325                 330                 335

tgc ttc ccg gcc aac ggc gca tgt ctg tgc gaa cac ggc ttc act ggg     1294
Cys Phe Pro Ala Asn Gly Ala Cys Leu Cys Glu His Gly Phe Thr Gly
340                 345                 350                 355

gac cgc tgc acg gat cgc ctc tgc ccc gac ggc ttc tac ggt ctc agc     1342
Asp Arg Cys Thr Asp Arg Leu Cys Pro Asp Gly Phe Tyr Gly Leu Ser
                360                 365                 370

tgc cag gcc ccc tgc acc tgc gac cgg gag cac agc ctc agc tgc cac     1390
Cys Gln Ala Pro Cys Thr Cys Asp Arg Glu His Ser Leu Ser Cys His
```

```
                375                 380                 385

ccg atg aac ggg gag tgc tcc tgc ctg ccg ggc tgg gcg ggc ctc cac     1438
Pro Met Asn Gly Glu Cys Ser Cys Leu Pro Gly Trp Ala Gly Leu His
        390                 395                 400

tgc aac gag agc tgc ccg cag gac acg cat ggg cca ggg tgc cag gag     1486
Cys Asn Glu Ser Cys Pro Gln Asp Thr His Gly Pro Gly Cys Gln Glu
    405                 410                 415

cac tgt ctc tgc ctg cac ggt ggc gtc tgc cag gct acc agc ggc ctc     1534
His Cys Leu Cys Leu His Gly Gly Val Cys Gln Ala Thr Ser Gly Leu
420                 425                 430                 435

tgt cag tgc gcg ccg ggt tac acg ggc cct cac tgt gct agt ctt tgt     1582
Cys Gln Cys Ala Pro Gly Tyr Thr Gly Pro His Cys Ala Ser Leu Cys
                440                 445                 450

cct cct gac acc tac ggt gtc aac tgt tct gca cgc tgc tca tgt gaa     1630
Pro Pro Asp Thr Tyr Gly Val Asn Cys Ser Ala Arg Cys Ser Cys Glu
            455                 460                 465

aat gcc atc gcc tgc tca ccc atc gac ggc gag tgc gtc tgc aag gaa     1678
Asn Ala Ile Ala Cys Ser Pro Ile Asp Gly Glu Cys Val Cys Lys Glu
        470                 475                 480

ggt tgg cag cgt ggt aac tgc tct gtg ccc tgc cca ccc gga acc tgg     1726
Gly Trp Gln Arg Gly Asn Cys Ser Val Pro Cys Pro Pro Gly Thr Trp
    485                 490                 495

ggc ttc agt tgc aat gcc agc tgc cag tgt gcc cat gag gca gtc tgc     1774
Gly Phe Ser Cys Asn Ala Ser Cys Gln Cys Ala His Glu Ala Val Cys
500                 505                 510                 515

agc ccc caa act gga gcc tgt acc tgc acc cct ggg tgg cat ggg gcc     1822
Ser Pro Gln Thr Gly Ala Cys Thr Cys Thr Pro Gly Trp His Gly Ala
                520                 525                 530

cac tgc cag ctg ccc tgt ccg aag ggg cag ttt gga gaa ggt tgt gcc     1870
His Cys Gln Leu Pro Cys Pro Lys Gly Gln Phe Gly Glu Gly Cys Ala
            535                 540                 545

agt cgc tgt gac tgt gac cac tct gat ggc tgt gac cct gtt cat gga     1918
Ser Arg Cys Asp Cys Asp His Ser Asp Gly Cys Asp Pro Val His Gly
        550                 555                 560

cgc tgt cag tgc cag gct ggc tgg atg ggt gcc cgc tgc cac ctg tcc     1966
Arg Cys Gln Cys Gln Ala Gly Trp Met Gly Ala Arg Cys His Leu Ser
    565                 570                 575

tgc cct gag ggc tta tgg gga gtc aac tgt agc aac acc tgc acc tgc     2014
Cys Pro Glu Gly Leu Trp Gly Val Asn Cys Ser Asn Thr Cys Thr Cys
580                 585                 590                 595

aag aat ggg ggc acc tgt ctc cct gag aat ggc aac tgc gtg tgt gca     2062
Lys Asn Gly Gly Thr Cys Leu Pro Glu Asn Gly Asn Cys Val Cys Ala
                600                 605                 610
```

```
ccc gga ttc cgg ggc ccc tcc tgc cag aga tcc tgt cag cct ggc cgc      2110
Pro Gly Phe Arg Gly Pro Ser Cys Gln Arg Ser Cys Gln Pro Gly Arg
            615                 620                 625

tat ggc aaa cgc tgt gtg ccc tgc aag tgc gct aac cac tcc ttc tgc      2158
Tyr Gly Lys Arg Cys Val Pro Cys Lys Cys Ala Asn His Ser Phe Cys
        630                 635                 640

cac ccc tcg aac ggg acc tgc tac tgc ctg gct ggc tgg aca ggc ccc      2206
His Pro Ser Asn Gly Thr Cys Tyr Cys Leu Ala Gly Trp Thr Gly Pro
    645                 650                 655

gac tgc tcc cag cca tgc cct cca gga cac tgg gga gaa aac tgt gcc      2254
Asp Cys Ser Gln Pro Cys Pro Pro Gly His Trp Gly Glu Asn Cys Ala
660                 665                 670                 675

cag acc tgc caa tgt cac cat ggt ggg acc tgc cat ccc cag gat ggg      2302
Gln Thr Cys Gln Cys His His Gly Gly Thr Cys His Pro Gln Asp Gly
                680                 685                 690

agc tgt atc tgc ccc cta ggc tgg act gga cac cac tgc tta gaa ggc      2350
Ser Cys Ile Cys Pro Leu Gly Trp Thr Gly His His Cys Leu Glu Gly
            695                 700                 705

tgc cct ctg ggg aca ttt ggt gct aac tgc tcc cag cca tgc cag tgt      2398
Cys Pro Leu Gly Thr Phe Gly Ala Asn Cys Ser Gln Pro Cys Gln Cys
        710                 715                 720

ggt cct gga gaa aag tgc cac cca gag act ggg gcc tgt gta tgt ccc      2446
Gly Pro Gly Glu Lys Cys His Pro Glu Thr Gly Ala Cys Val Cys Pro
    725                 730                 735

cca ggg cac agt ggt gca cct tgc agg att gga atc cag gag ccc ttt      2494
Pro Gly His Ser Gly Ala Pro Cys Arg Ile Gly Ile Gln Glu Pro Phe
740                 745                 750                 755

act gtg atg ccg acc act cca gta gcg tat aac tcg ctg ggt gca gtg      2542
Thr Val Met Pro Thr Thr Pro Val Ala Tyr Asn Ser Leu Gly Ala Val
                760                 765                 770

att ggc att gca gtg ctg ggg tcc ctt gtg gta gcc ctg gtg gca ctg      2590
Ile Gly Ile Ala Val Leu Gly Ser Leu Val Val Ala Leu Val Ala Leu
            775                 780                 785

ttc att ggc tat cgg cac tgg caa aaa ggc aag gag cac cac cac ctg      2638
Phe Ile Gly Tyr Arg His Trp Gln Lys Gly Lys Glu His His His Leu
        790                 795                 800

gct gtg gct tac agc agc ggg cgc ctg gac ggc tcc gag tat gtc atg      2686
Ala Val Ala Tyr Ser Ser Gly Arg Leu Asp Gly Ser Glu Tyr Val Met
    805                 810                 815

cca gat gtc cct ccg agc tac agt cac tac tac tcc aac ccc agc tac      2734
Pro Asp Val Pro Pro Ser Tyr Ser His Tyr Tyr Ser Asn Pro Ser Tyr
820                 825                 830                 835

cac acc ctg tcg cag tgc tcc cca aac ccc cca ccc cct aac aag gtt      2782
His Thr Leu Ser Gln Cys Ser Pro Asn Pro Pro Pro Pro Asn Lys Val
```

```
                840                 845                 850

cca ggc ccg ctc ttt gcc agc ctg cag aac cct gag cgg cca ggt ggg     2830
Pro Gly Pro Leu Phe Ala Ser Leu Gln Asn Pro Glu Arg Pro Gly Gly
            855                 860                 865

gcc caa ggg cat gat aac cac acc acc ctg cct gct gac tgg aag cac     2878
Ala Gln Gly His Asp Asn His Thr Thr Leu Pro Ala Asp Trp Lys His
        870                 875                 880

cgc cgg gag ccc cct cca ggg cct ctg gac agg ggg agc agc cgc ctg     2926
Arg Arg Glu Pro Pro Pro Gly Pro Leu Asp Arg Gly Ser Ser Arg Leu
    885                 890                 895

gac cga agc tac agc tat agc tac agc aat ggc cca ggc cca ttc tac     2974
Asp Arg Ser Tyr Ser Tyr Ser Tyr Ser Asn Gly Pro Gly Pro Phe Tyr
900                 905                 910                 915

gat aaa ggg ctc atc tct gaa gag gag ctc ggg gcc agt gtg gct tcc     3022
Asp Lys Gly Leu Ile Ser Glu Glu Glu Leu Gly Ala Ser Val Ala Ser
                920                 925                 930

ctg agc agt gag aac cca tat gcc acc atc cgg gac ctg ccc agc ttg     3070
Leu Ser Ser Glu Asn Pro Tyr Ala Thr Ile Arg Asp Leu Pro Ser Leu
            935                 940                 945

cca ggg ggc ccc cgg gag agc agc tac atg gag atg aaa ggc cct ccc     3118
Pro Gly Gly Pro Arg Glu Ser Ser Tyr Met Glu Met Lys Gly Pro Pro
        950                 955                 960

tca gga tct gcc ccc agg cag cct cct cag ttt tgg gac agc cag agg     3166
Ser Gly Ser Ala Pro Arg Gln Pro Pro Gln Phe Trp Asp Ser Gln Arg
    965                 970                 975

cgg cgg caa ccc cag cca cag aga gac agt ggc acc tac gag cag ccc     3214
Arg Arg Gln Pro Gln Pro Gln Arg Asp Ser Gly Thr Tyr Glu Gln Pro
980                 985                 990                 995

agc ccc ctg atc cat gac cga gac tct gtg ggc tcc cag ccc cct ctg     3262
Ser Pro Leu Ile His Asp Arg Asp Ser Val Gly Ser Gln Pro Pro Leu
                1000                1005                1010

cct ccg ggc cta ccc ccc ggc cac tat gac tca ccc aag aac agc cac     3310
Pro Pro Gly Leu Pro Pro Gly His Tyr Asp Ser Pro Lys Asn Ser His
            1015                1020                1025

atc cct gga cat tat gac ttg cct cca gta cgg cat ccc cca tca cct     3358
Ile Pro Gly His Tyr Asp Leu Pro Pro Val Arg His Pro Pro Ser Pro
        1030                1035                1040

cca ctt cga cgc cag gac cgt tgaggagcca ggatggtatg gcagaggcca        3409
Pro Leu Arg Arg Gln Asp Arg
    1045                1050

gcacacctgg ctgttgctgc tcaaggctgg ggacagagcc tagtgtaccc ctgccaggag   3469
cagggagtgg accggcaggc tgtgaacatg aacaacgctt aacagagcaa gtgatgggag   3529
ccttgttcct gggttctacc atgggagacg ctgatcagca ggatgcctgg ctccctttcc   3589
caacccactg ctcccaaggc ctccagggcc ctgtgtacat aaactggtgg gttggaagtt   3649
```

gctgggtaac tctgatttca gacatgcgtg tggggtacct tttctgtgca tgctcagcct 3709
gggctctgtg cgtgtgtgtg tttctgtgat tttagaaggg taccaggcag gttctgtcct 3769
agggcactta ccatttagta gggagatgga accaacccaa ttaactctag caatagcctc 3829
ctaactggcc tcctccattg attcagtgaa ccttccaatg catggctcat aatttcaaaa 3889
tacaggctgg ttagttactc cctacctgaa agccttcata ggtgcctctt tgctcttctg 3949
ccagtatcaa aacttttgaa ggccttaaag gccctgcttt gcctggccca tctgtctctc 4009
cagcctcacc ttgaactgtg ttcctgtcac tgcacgccag tcacaccggc ctctaggtcc 4069
tcctgtaggc cactcttctt tctggcacag ggacctgcac acctggagtg cccttcctcc 4129
cccactcgcc tgttcacccc tgcttttcct ttacacctcc tcctcaggga agtgcccacc 4189
ctccgtacat ctttcacagc cctgattgca gctgtgttca ctcaccaggt acctgcagaa 4249
ggcctacagg gtgccaggca cttctttaat gggttctttc tttatgtgat tatttgatta 4309
atctctgcct cccccactag actgtaagct ccctgaaggc aagaatcctg tgcttatgct 4369
caatattagc tctcccttgg cacagagtag gcactcaaca aatgctcccc aaaaggctga 4429
gtggctgact gaattaagta ccagtgacat gcagtaactg ctaagataga tgagccatct 4489
gtatgctctg acagttacag actgaataag ttggagactt ccctaaaggg tggcatttcc 4549
ccagggtaac aacgcagagc tcaggtgtgg gaaggtgcca ggggcagggg tgcagagggg 4609
ctgaggctga ggggggtgca gaggctggag aaaggataac aggagagagt atacaggcat 4669
gccttgattt attgcacttc acaggtagca gaatttttaa agaaattgaa ggttttggga 4729
catatatgtg acagcaatag gttaagaaaa gcaaagcaga gaaattgaag atttgtgtca 4789
acactgcttt aagcaaatct gttggcacca tttttccaat agcatgtgcc cattttgggt 4849
ctctacattg cattttggta attgcttgca atatttcaag cattttcatt gttattatat 4909
gtgttatagt gatctgtgat cagtgatctt tgatatatta ttgtaattgt ttcggggcgc 4969
catgaaccgc acccatataa cacggtaaac ttaatcagca aaaaaaaaaa aaaaaaaagg 5029
gcggccg 5036

<210> 14
<211> 1050
<212> PRT
<213> Homo sapiens

<400> 14

```
Met Ser Pro Pro Leu Cys Pro Leu Leu Leu Leu Ala Val Gly Leu Arg
 1               5                  10                  15
Leu Ala Gly Thr Leu Asn Pro Ser Asp Pro Asn Thr Cys Ser Phe Trp
            20                  25                  30
Glu Ser Phe Thr Thr Thr Thr Lys Glu Ser His Ser Arg Pro Phe Ser
        35                  40                  45
Leu Leu Pro Ser Glu Pro Cys Glu Arg Pro Trp Glu Gly Pro His Thr
    50                  55                  60
Cys Pro Ser Pro Gln Thr Gln Arg Lys Leu Leu Ala Ser Arg Asp Ser
65                  70                  75                  80
Phe Cys Met Val Cys Val Gly Ala Gly Val Gln Trp Arg Asp Arg Ser
                85                  90                  95
Ala Leu Gln Pro Gln Thr Gly Asn Ala Leu Ser Met Arg Pro Gln Pro
            100                 105                 110
Arg Val Leu Ser Gly Ala Pro Ser Leu Ala Ser Pro Gly His Thr Val
        115                 120                 125
Val Val Lys Thr Asp His Arg Gln Arg Leu Gln Cys Cys His Gly Phe
    130                 135                 140
Tyr Glu Ser Arg Gly Phe Cys Val Pro Leu Cys Ala Gln Glu Cys Val
145                 150                 155                 160
His Gly Arg Cys Val Ala Pro Asn Gln Cys Gln Cys Val Pro Gly Trp
                165                 170                 175
Arg Gly Asp Asp Cys Ser Ser Ala Pro Asn Cys Leu Gln Pro Cys Thr
            180                 185                 190
Pro Gly Tyr Tyr Gly Pro Ala Cys Gln Phe Arg Cys Gln Cys His Gly
        195                 200                 205
Ala Pro Cys Asp Pro Gln Thr Gly Ala Cys Phe Cys Pro Ala Glu Arg
```

```
    210                 215                 220
Thr Gly Pro Ser Cys Asp Val Ser Cys Ser Gln Gly Thr Ser Gly Phe
225                 230                 235                 240
Phe Cys Pro Ser Thr His Pro Cys Gln Asn Gly Gly Val Phe Gln Thr
                245                 250                 255
Pro Gln Gly Ser Cys Ser Cys Pro Pro Gly Trp Met Gly Thr Ile Cys
            260                 265                 270
Ser Leu Pro Cys Pro Glu Gly Phe His Gly Pro Asn Cys Ser Gln Glu
        275                 280                 285
Cys Arg Cys His Asn Gly Gly Leu Cys Asp Arg Phe Thr Gly Gln Cys
    290                 295                 300
Arg Cys Ala Pro Gly Tyr Thr Gly Asp Arg Cys Arg Glu Glu Cys Pro
305                 310                 315                 320
Val Gly Arg Phe Gly Gln Asp Cys Ala Glu Thr Cys Asp Cys Ala Pro
                325                 330                 335
Asp Ala Arg Cys Phe Pro Ala Asn Gly Ala Cys Leu Cys Glu His Gly
            340                 345                 350
Phe Thr Gly Asp Arg Cys Thr Asp Arg Leu Cys Pro Asp Gly Phe Tyr
        355                 360                 365
Gly Leu Ser Cys Gln Ala Pro Cys Thr Cys Asp Arg Glu His Ser Leu
    370                 375                 380
Ser Cys His Pro Met Asn Gly Glu Cys Ser Cys Leu Pro Gly Trp Ala
385                 390                 395                 400
Gly Leu His Cys Asn Glu Ser Cys Pro Gln Asp Thr His Gly Pro Gly
                405                 410                 415
Cys Gln Glu His Cys Leu Cys Leu His Gly Gly Val Cys Gln Ala Thr
            420                 425                 430
Ser Gly Leu Cys Gln Cys Ala Pro Gly Tyr Thr Gly Pro His Cys Ala
        435                 440                 445
Ser Leu Cys Pro Pro Asp Thr Tyr Gly Val Asn Cys Ser Ala Arg Cys
    450                 455                 460
Ser Cys Glu Asn Ala Ile Ala Cys Ser Pro Ile Asp Gly Glu Cys Val
465                 470                 475                 480
Cys Lys Glu Gly Trp Gln Arg Gly Asn Cys Ser Val Pro Cys Pro Pro
                485                 490                 495
Gly Thr Trp Gly Phe Ser Cys Asn Ala Ser Cys Gln Cys Ala His Glu
            500                 505                 510
Ala Val Cys Ser Pro Gln Thr Gly Ala Cys Thr Cys Thr Pro Gly Trp
        515                 520                 525
His Gly Ala His Cys Gln Leu Pro Cys Pro Lys Gly Gln Phe Gly Glu
    530                 535                 540
Gly Cys Ala Ser Arg Cys Asp Cys Asp His Ser Asp Gly Cys Asp Pro
545                 550                 555                 560
Val His Gly Arg Cys Gln Cys Gln Ala Gly Trp Met Gly Ala Arg Cys
                565                 570                 575
His Leu Ser Cys Pro Glu Gly Leu Trp Gly Val Asn Cys Ser Asn Thr
            580                 585                 590
Cys Thr Cys Lys Asn Gly Gly Thr Cys Leu Pro Glu Asn Gly Asn Cys
        595                 600                 605
Val Cys Ala Pro Gly Phe Arg Gly Pro Ser Cys Gln Arg Ser Cys Gln
    610                 615                 620
Pro Gly Arg Tyr Gly Lys Arg Cys Val Pro Cys Lys Cys Ala Asn His
625                 630                 635                 640
Ser Phe Cys His Pro Ser Asn Gly Thr Cys Tyr Cys Leu Ala Gly Trp
                645                 650                 655
Thr Gly Pro Asp Cys Ser Gln Pro Cys Pro Pro Gly His Trp Gly Glu
            660                 665                 670
Asn Cys Ala Gln Thr Cys Gln Cys His His Gly Gly Thr Cys His Pro
```

```
        675                 680                 685
Gln Asp Gly Ser Cys Ile Cys Pro Leu Gly Trp Thr Gly His His Cys
    690                 695                 700
Leu Glu Gly Cys Pro Leu Gly Thr Phe Gly Ala Asn Cys Ser Gln Pro
705                 710                 715                 720
Cys Gln Cys Gly Pro Gly Glu Lys Cys His Pro Glu Thr Gly Ala Cys
                725                 730                 735
Val Cys Pro Pro Gly His Ser Gly Ala Pro Cys Arg Ile Gly Ile Gln
            740                 745                 750
Glu Pro Phe Thr Val Met Pro Thr Thr Pro Val Ala Tyr Asn Ser Leu
        755                 760                 765
Gly Ala Val Ile Gly Ile Ala Val Leu Gly Ser Leu Val Val Ala Leu
    770                 775                 780
Val Ala Leu Phe Ile Gly Tyr Arg His Trp Gln Lys Gly Lys Glu His
785                 790                 795                 800
His His Leu Ala Val Ala Tyr Ser Ser Gly Arg Leu Asp Gly Ser Glu
                805                 810                 815
Tyr Val Met Pro Asp Val Pro Pro Ser Tyr Ser His Tyr Tyr Ser Asn
            820                 825                 830
Pro Ser Tyr His Thr Leu Ser Gln Cys Ser Pro Asn Pro Pro Pro Pro
        835                 840                 845
Asn Lys Val Pro Gly Pro Leu Phe Ala Ser Leu Gln Asn Pro Glu Arg
    850                 855                 860
Pro Gly Gly Ala Gln Gly His Asp Asn His Thr Thr Leu Pro Ala Asp
865                 870                 875                 880
Trp Lys His Arg Arg Glu Pro Pro Pro Gly Pro Leu Asp Arg Gly Ser
                885                 890                 895
Ser Arg Leu Asp Arg Ser Tyr Ser Tyr Ser Tyr Ser Asn Gly Pro Gly
            900                 905                 910
Pro Phe Tyr Asp Lys Gly Leu Ile Ser Glu Glu Glu Leu Gly Ala Ser
        915                 920                 925
Val Ala Ser Leu Ser Ser Glu Asn Pro Tyr Ala Thr Ile Arg Asp Leu
    930                 935                 940
Pro Ser Leu Pro Gly Gly Pro Arg Glu Ser Ser Tyr Met Glu Met Lys
945                 950                 955                 960
Gly Pro Pro Ser Gly Ser Ala Pro Arg Gln Pro Pro Gln Phe Trp Asp
                965                 970                 975
Ser Gln Arg Arg Arg Gln Pro Gln Pro Gln Arg Asp Ser Gly Thr Tyr
            980                 985                 990
Glu Gln Pro Ser Pro Leu Ile His Asp Arg Asp Ser Val Gly Ser Gln
        995                 1000                1005
Pro Pro Leu Pro Pro Gly Leu Pro Pro Gly His Tyr Asp Ser Pro Lys
    1010                1015                1020
Asn Ser His Ile Pro Gly His Tyr Asp Leu Pro Pro Val Arg His Pro
1025                1030                1035                1040
Pro Ser Pro Pro Leu Arg Arg Gln Asp Arg
                1045                1050
```

<210> 15
<211> 3150
<212> DNA
<213> Homo sapiens

<400> 15

```
atgtcaccgc ctctgtgtcc cctccttctc ctggctgtgg gcctgcggct ggctggaact      60
ctcaacccca gtgatcccaa tacctgcagc ttctgggaaa gcttcactac caccaccaag     120
gagtcccact cccgcccctt cagcctgctc ccctcagagc cctgcgagcg gccctgggag     180
ggcccccata cttgccccag cccacaaact cagaggaaac tcctggcttc tagggattca     240
```

ttctgcatgg tctgtgtcgg ggctggagtg cagtggcgag atcgtagtgc actgcaacct 300
caaacaggga atgcgctttc tatgcgccct cagcccagag tgttgagtgg tgccccttcc 360
ctggcctccc ctggccacac tgtggtggtg aagacggacc accgccagcg cctgcagtgc 420
tgccatggct tctatgagag caggggttc tgtgtcccgc tctgtgccca ggagtgtgtc 480
catggccgtt gtgtggcacc caatcagtgc caatgtgtgc caggctggcg gggcgacgac 540
tgttccagtg ccccgaactg ccttcagccc tgtacccctg gctactatgg ccctgcctgc 600
cagttccgct gccagtgcca tggggcaccc tgcgatcccc agactggagc ctgcttctgc 660
cccgcagaga gaactgggcc cagctgtgac gtgtcctgtt cccagggcac ttctggcttc 720
ttctgcccca gcacccatcc ttgccaaaat ggaggtgtct tccaaacccc acagggctcc 780
tgcagctgcc cccctggctg gatgggcacc atctgctccc tgccctgccc agagggcttt 840
cacggaccca actgctccca ggaatgtcgc tgccacaacg gcggcctctg tgaccgattc 900
actgggcagt gccgctgcgc tccgggttac actggggatc ggtgccggga ggagtgcccg 960
gtgggccgct ttgggcagga ctgtgctgag acgtgcgact gcgccccgga cgcccgttgc 1020
ttcccggcca acggcgcatg tctgtgcgaa cacggcttca ctggggaccg ctgcacggat 1080
cgcctctgcc ccgacggctt ctacggtctc agctgccagg ccccctgcac ctgcgaccgg 1140
gagcacagcc tcagctgcca cccgatgaac ggggagtgct cctgcctgcc gggctgggcg 1200
ggcctccact gcaacgagag ctgcccgcag gacacgcatg ggccagggtg ccaggagcac 1260
tgtctctgcc tgcacggtgg cgtctgccag gctaccagcg gcctctgtca gtgcgcgccg 1320
ggttacacgg gccctcactg tgctagtctt tgtcctcctg acacctacgg tgtcaactgt 1380
tctgcacgct gctcatgtga aaatgccatc gcctgctcac ccatcgacgg cgagtgcgtc 1440
tgcaaggaag gttggcagcg tggtaactgc tctgtgccct gcccacccgg aacctggggc 1500
ttcagttgca atgccagctg ccagtgtgcc catgaggcag tctgcagccc ccaaactgga 1560
gcctgtacct gcacccctgg gtggcatggg gcccactgcc agctgccctg tccgaagggg 1620
cagtttggag aaggttgtgc cagtcgctgt gactgtgacc actctgatgg ctgtgaccct 1680
gttcatggac gctgtcagtg ccaggctggc tggatgggtg cccgctgcca cctgtcctgc 1740
cctgagggct tatggggagt caactgtagc aacacctgca cctgcaagaa tgggggcacc 1800
tgtctccctg agaatggcaa ctgcgtgtgt gcacccggat tccggggccc ctcctgccag 1860
agatcctgtc agcctggccg ctatggcaaa cgctgtgtgc cctgcaagtg cgctaaccac 1920
tccttctgcc acccctcgaa cgggacctgc tactgcctgg ctggctggac aggccccgac 1980
tgctcccagc catgccctcc aggacactgg ggagaaaact gtgcccagac ctgccaatgt 2040
caccatggtg ggacctgcca tccccaggat gggagctgta tctgcccct aggctggact 2100
ggacaccact gcttagaagg ctgccctctg gggacatttg gtgctaactg ctcccagcca 2160
tgccagtgtg gtcctggaga aaagtgccac ccagagactg gggcctgtgt atgtcccca 2220
gggcacagtg gtgcaccttg caggattgga atccaggagc cctttactgt gatgccgacc 2280
actccagtag cgtataactc gctgggtgca gtgattggca ttgcagtgct ggggtccctt 2340
gtggtagccc tggtggcact gttcattggc tatcggcact ggcaaaaagg caaggagcac 2400
caccacctgg ctgtggctta cagcagcggg cgcctggacg gctccgagta tgtcatgcca 2460
gatgtccctc cgagctacag tcactactac tccaacccca gctaccacac cctgtcgcag 2520
tgctccccaa acccccccacc ccctaacaag gttccaggcc cgctctttgc cagcctgcag 2580
aaccctgagc ggccaggtgg ggcccaaggg catgataacc acaccgccct gcctgctgac 2640
tggaagcacc gccgggagcc ccctccaggg cctctggaca gggggagcag ccgcctggac 2700
cgaagctaca gctatagcta cagcaatggc ccaggcccat tctacgataa agggctcatc 2760
tctgaagagg agctcggggc cagtgtggct tccctgagca gtgagaaccc atatgccacc 2820
atccgggacc tgcccagctt gccagggggc ccccgggaga gcagctacat ggagatgaaa 2880
ggccctccct caggatctgc ccccaggcag cctcctcagt tttgggacag ccagaggcgg 2940
cggcaacccc agccacagag agacagtggc acctacgagc agcccagccc cctgatccat 3000
gaccgagact ctgtgggctc ccagcccccct ctgcctccgg gcctaccccc cggccactat 3060
gactcaccca agaacagcca catccctgga cattatgact tgcctccagt acggcatccc 3120
ccatcacctc cacttcgacg ccaggaccgt 3150

<210> 16
<211> 2569
<212> DNA
<213> Mus musculus

<220>
<221> CDS
<222> (2)...(1492)

<400> 16

```
g tcg acc cac gcg tcc ggt gac cct gtt cat gga cag tgc cga tgt cag      49
  Ser Thr His Ala Ser Gly Asp Pro Val His Gly Gln Cys Arg Cys Gln
   1               5                  10                  15

gct ggt tgg atg ggc aca cgc tgc cac ctg cct tgc ccg gag ggc ttt      97
Ala Gly Trp Met Gly Thr Arg Cys His Leu Pro Cys Pro Glu Gly Phe
            20                  25                  30

tgg gga gcc aac tgc agt aac acc tgt acc tgc aag aat ggt ggt acc     145
Trp Gly Ala Asn Cys Ser Asn Thr Cys Thr Cys Lys Asn Gly Gly Thr
        35                  40                  45

tgt gtg tct gag aat ggc aac tgc gtg tgc gca cca ggg ttc cga ggc     193
Cys Val Ser Glu Asn Gly Asn Cys Val Cys Ala Pro Gly Phe Arg Gly
    50                  55                  60

ccc tcc tgc cag agg ccc tgc ccg cct ggt cgc tat ggc aaa cgc tgt     241
Pro Ser Cys Gln Arg Pro Cys Pro Pro Gly Arg Tyr Gly Lys Arg Cys
65                  70                  75                  80

gtg caa tgc aag tgt aac aac aac cat tct tcc tgc cac cca tcg gac     289
Val Gln Cys Lys Cys Asn Asn Asn His Ser Ser Cys His Pro Ser Asp
                85                  90                  95

ggg acc tgc tcc tgc ctg gcg ggc tgg aca ggc cct gac tgc tcc gag     337
Gly Thr Cys Ser Cys Leu Ala Gly Trp Thr Gly Pro Asp Cys Ser Glu
            100                 105                 110

gca tgt ccc cca ggc cac tgg gga ctc aaa tgc tcc caa ctc tgc cag     385
Ala Cys Pro Pro Gly His Trp Gly Leu Lys Cys Ser Gln Leu Cys Gln
        115                 120                 125

tgt cat cat ggt ggg acc tgc cac ccc cag gat ggg agc tgt atc tgc     433
Cys His His Gly Gly Thr Cys His Pro Gln Asp Gly Ser Cys Ile Cys
    130                 135                 140

acg cca ggc tgg act gga ccc aac tgc ttg gaa ggc tgc cca cca aga     481
Thr Pro Gly Trp Thr Gly Pro Asn Cys Leu Glu Gly Cys Pro Pro Arg
145                 150                 155                 160

atg ttt ggt gtc aac tgc tcc cag cta tgt cag tgt gat ctc gga gag     529
Met Phe Gly Val Asn Cys Ser Gln Leu Cys Gln Cys Asp Leu Gly Glu
                165                 170                 175

atg tgc cac cca gag act ggg gct tgt gtc tgt ccc cca gga cac agt     577
Met Cys His Pro Glu Thr Gly Ala Cys Val Cys Pro Pro Gly His Ser
            180                 185                 190

ggt gca gac tgc aaa atg gga agc cag gag tcc ttc acc ata atg ccc     625
Gly Ala Asp Cys Lys Met Gly Ser Gln Glu Ser Phe Thr Ile Met Pro
        195                 200                 205

acc tct ccc gtg acc cat aac tca ctg ggt gca gtg att ggc att gca     673
Thr Ser Pro Val Thr His Asn Ser Leu Gly Ala Val Ile Gly Ile Ala
    210                 215                 220

gta ctg gga acc ctc gtg gtg gcc ctg ata gca ctg ttc att ggc tac     721
```

```
Val Leu Gly Thr Leu Val Val Ala Leu Ile Ala Leu Phe Ile Gly Tyr
225                 230                 235                 240

cgc cag tgg caa aag ggc aag gaa cat gag cac ttg gca gtg gct tac      769
Arg Gln Trp Gln Lys Gly Lys Glu His Glu His Leu Ala Val Ala Tyr
                245                 250                 255

agc act ggg cgg ctg gat ggc tct gat tac gtc atg cca gat gtc tct      817
Ser Thr Gly Arg Leu Asp Gly Ser Asp Tyr Val Met Pro Asp Val Ser
            260                 265                 270

ccg agc tat agt cac tac tac tcc aac ccc agc tac cac aca ctg tct      865
Pro Ser Tyr Ser His Tyr Tyr Ser Asn Pro Ser Tyr His Thr Leu Ser
        275                 280                 285

cag tgt tct cct aac ccc ccg ccc cct aac aag gtc cca ggc agt cag      913
Gln Cys Ser Pro Asn Pro Pro Pro Pro Asn Lys Val Pro Gly Ser Gln
    290                 295                 300

ctc ttt gtc agc tct cag gcc cct gag cgg cca agc aga gcc cac ggg      961
Leu Phe Val Ser Ser Gln Ala Pro Glu Arg Pro Ser Arg Ala His Gly
305                 310                 315                 320

cgt gag aac cat acc aca ctg ccc gct gac tgg aag cac cgc cgg gag     1009
Arg Glu Asn His Thr Thr Leu Pro Ala Asp Trp Lys His Arg Arg Glu
                325                 330                 335

ccc cat gac aga ggc gcc agc cac ctg gac cga agc tat agc tgt agc     1057
Pro His Asp Arg Gly Ala Ser His Leu Asp Arg Ser Tyr Ser Cys Ser
            340                 345                 350

tat agc cac agg aat ggc cca gga cca ttc tgt cat aaa ggt ccc atc     1105
Tyr Ser His Arg Asn Gly Pro Gly Pro Phe Cys His Lys Gly Pro Ile
        355                 360                 365

tct gaa gag gga cta ggg gca agc gtt atg tcc ctg agc agt gag aac     1153
Ser Glu Glu Gly Leu Gly Ala Ser Val Met Ser Leu Ser Ser Glu Asn
    370                 375                 380

ccc tat gct acc atc cga gac ctg ccc agc ctg cct ggg gaa ccc cga     1201
Pro Tyr Ala Thr Ile Arg Asp Leu Pro Ser Leu Pro Gly Glu Pro Arg
385                 390                 395                 400

gaa agt ggc tat gtg gag atg aaa gga cct cca tca gtg tcc cct ccc     1249
Glu Ser Gly Tyr Val Glu Met Lys Gly Pro Pro Ser Val Ser Pro Pro
                405                 410                 415

agg cag tct ctt cat ctc cgg gac agg cag cag cgg caa ctg cag cca     1297
Arg Gln Ser Leu His Leu Arg Asp Arg Gln Gln Arg Gln Leu Gln Pro
            420                 425                 430

cag agg gac agc ggc acc tat gag cag ccc agc ccc ttg agc cat aat     1345
Gln Arg Asp Ser Gly Thr Tyr Glu Gln Pro Ser Pro Leu Ser His Asn
        435                 440                 445

gaa gag tct ttg ggc tcc acg ccc ccg ctt cct cca ggc ctg cct cct     1393
Glu Glu Ser Leu Gly Ser Thr Pro Pro Leu Pro Pro Gly Leu Pro Pro
    450                 455                 460
```

```
ggt cac tac gac tcc ccc aag aac agc cat atc cct gga cac tat gac     1441
Gly His Tyr Asp Ser Pro Lys Asn Ser His Ile Pro Gly His Tyr Asp
465                 470                 475                 480

ttg cct cca gta cgg cat cct cca tcc cct cca tcc cgg cgc cag gac     1489
Leu Pro Pro Val Arg His Pro Pro Ser Pro Pro Ser Arg Arg Gln Asp
                485                 490                 495

cgc tgaagagccg gcatggtatg ggagcgtgcc tatgtacctt gccaggagca          1542
Arg


gggactggac cagcaggcca cgaacagaaa cacttggtga agtgaacaga gacggactgt   1602
ggccctgtgc ttccaccgag ggagacacta gttgacaaag tgtctaaccc tcttttccaa   1662
cccactgctc aagtccctgt ggacataagc tggtgggcag aatgttgttg tacaagtgtg   1722
attttagatc gatttttttt taaagtatgt gttgggtacc ttttctgtgt gtatgctcag   1782
gcaggctgtg tgtgtctcta gttggcttta gagggagtca ggtataggtt ctgccttctg   1842
cactttccat cttatctagt agtcagcttc caagcttaac tagttagagc tccaccagca   1902
gcaggcccta actacctgcc tgcccttcac ccagtaatcc tccatgtctt tgctcagagg   1962
attgctcccc gactctggtg ttgtcctcct ggtacgcctt gacggtcctg cagtctccct   2022
ttcccgtctt gcttcattct ttcccagaat gaaggctgtc tgccacccta cttcccagcc   2082
caggaattgg cacatctaag ttcagccttc ctaagttacc cgttgagtcc tgcttgccct   2142
tcacatattc cacagaacac ccaccccaca tctgcttcat agctactctc ttctccacgt   2202
acccacagaa ggcagaagtg gtaccaggca agaagatggg attgttgcat tttgttttgt   2262
ttttgagact ctgtctcact atgtagtcct ggctggcctg gaactcaaga gctctgcctg   2322
cctctgcctc ttgagtgctg ggtttaacgg ctcagggtca catgcacagc tcaagctgca   2382
ctccgatgtg ctttcccctg ttgctagatt agcgtctgcc tccccctagt ggagaggctg   2442
atcgccagct ctctgatgca ggactctggt gtttaggctc actcactatt ggtttccttg   2502
gcacagggta gtcactcaat aaatgttcct ctaaaagctg aaaaaaaaaa aaaaaaaggg   2562
cggccgc                                                            2569
```

<210> 17
<211> 497
<212> PRT
<213> Mus musculus

<400> 17

```
Ser Thr His Ala Ser Gly Asp Pro Val His Gly Gln Cys Arg Cys Gln
 1               5                  10                  15
Ala Gly Trp Met Gly Thr Arg Cys His Leu Pro Cys Pro Glu Gly Phe
            20                  25                  30
Trp Gly Ala Asn Cys Ser Asn Thr Cys Thr Cys Lys Asn Gly Gly Thr
        35                  40                  45
Cys Val Ser Glu Asn Gly Asn Cys Val Cys Ala Pro Gly Phe Arg Gly
    50                  55                  60
Pro Ser Cys Gln Arg Pro Cys Pro Pro Gly Arg Tyr Gly Lys Arg Cys
65                  70                  75                  80
Val Gln Cys Lys Cys Asn Asn Asn His Ser Ser Cys His Pro Ser Asp
                85                  90                  95
Gly Thr Cys Ser Cys Leu Ala Gly Trp Thr Gly Pro Asp Cys Ser Glu
            100                 105                 110
Ala Cys Pro Pro Gly His Trp Gly Leu Lys Cys Ser Gln Leu Cys Gln
        115                 120                 125
Cys His His Gly Gly Thr Cys His Pro Gln Asp Gly Ser Cys Ile Cys
    130                 135                 140
Thr Pro Gly Trp Thr Gly Pro Asn Cys Leu Glu Gly Cys Pro Pro Arg
145                 150                 155                 160
```

```
Met Phe Gly Val Asn Cys Ser Gln Leu Cys Gln Cys Asp Leu Gly Glu
                165                 170                 175
Met Cys His Pro Glu Thr Gly Ala Cys Val Cys Pro Pro Gly His Ser
            180                 185                 190
Gly Ala Asp Cys Lys Met Gly Ser Gln Glu Ser Phe Thr Ile Met Pro
        195                 200                 205
Thr Ser Pro Val Thr His Asn Ser Leu Gly Ala Val Ile Gly Ile Ala
    210                 215                 220
Val Leu Gly Thr Leu Val Val Ala Leu Ile Ala Leu Phe Ile Gly Tyr
225                 230                 235                 240
Arg Gln Trp Gln Lys Gly Lys Glu His Glu His Leu Ala Val Ala Tyr
                245                 250                 255
Ser Thr Gly Arg Leu Asp Gly Ser Asp Tyr Val Met Pro Asp Val Ser
            260                 265                 270
Pro Ser Tyr Ser His Tyr Tyr Ser Asn Pro Ser Tyr His Thr Leu Ser
        275                 280                 285
Gln Cys Ser Pro Asn Pro Pro Pro Pro Asn Lys Val Pro Gly Ser Gln
    290                 295                 300
Leu Phe Val Ser Ser Gln Ala Pro Glu Arg Pro Ser Arg Ala His Gly
305                 310                 315                 320
Arg Glu Asn His Thr Thr Leu Pro Ala Asp Trp Lys His Arg Arg Glu
                325                 330                 335
Pro His Asp Arg Gly Ala Ser His Leu Asp Arg Ser Tyr Ser Cys Ser
            340                 345                 350
Tyr Ser His Arg Asn Gly Pro Gly Pro Phe Cys His Lys Gly Pro Ile
        355                 360                 365
Ser Glu Glu Gly Leu Gly Ala Ser Val Met Ser Leu Ser Ser Glu Asn
    370                 375                 380
Pro Tyr Ala Thr Ile Arg Asp Leu Pro Ser Leu Pro Gly Glu Pro Arg
385                 390                 395                 400
Glu Ser Gly Tyr Val Glu Met Lys Gly Pro Pro Ser Val Ser Pro Pro
                405                 410                 415
Arg Gln Ser Leu His Leu Arg Asp Arg Gln Gln Arg Gln Leu Gln Pro
            420                 425                 430
Gln Arg Asp Ser Gly Thr Tyr Glu Gln Pro Ser Pro Leu Ser His Asn
        435                 440                 445
Glu Glu Ser Leu Gly Ser Thr Pro Pro Leu Pro Pro Gly Leu Pro Pro
    450                 455                 460
Gly His Tyr Asp Ser Pro Lys Asn Ser His Ile Pro Gly His Tyr Asp
465                 470                 475                 480
Leu Pro Pro Val Arg His Pro Pro Ser Pro Pro Ser Arg Arg Gln Asp
                485                 490                 495
Arg
```

<210> 18
<211> 1491
<212> DNA
<213> Mus musculus

<400> 18

```
tcgacccacg cgtccggtga ccctgttcat ggacagtgcc gatgtcaggc tggttggatg      60
ggcacacgct gccacctgcc ttgcccggag ggcttttggg gagccaactg cagtaacacc     120
tgtacctgca agaatggtgg tacctgtgtg tctgagaatg gcaactgcgt gtgcgcacca     180
gggttccgag gcccctcctg ccagaggccc tgcccgcctg gtcgctatgg caaacgctgt     240
gtgcaatgca agtgtaacaa caaccattct tcctgccacc catcggacgg gacctgctcc     300
tgcctggcgg gctggacagg ccctgactgc tccgaggcat gtcccccagg ccactgggga     360
ctcaaatgct cccaactctg ccagtgtcat catggtggga cctgccaccc ccaggatggg     420
```

```
agctgtatct gcacgccagg ctggactgga cccaactgct tggaaggctg cccaccaaga     480
atgtttggtg tcaactgctc ccagctatgt cagtgtgatc tcggagagat gtgccaccca     540
gagactgggg cttgtgtctg tcccccagga cacagtggtg cagactgcaa aatgggaagc     600
caggagtcct tcaccataat gcccacctct cccgtgaccc ataactcact gggtgcagtg     660
attggcattg cagtactggg aaccctcgtg gtggccctga tagcactgtt cattggctac     720
cgccagtggc aaaagggcaa ggaacatgag cacttggcag tggcttacag cactgggcgg     780
ctggatggct ctgattacgt catgccagat gtctctccga gctatagtca ctactactcc     840
aaccccagct accacacact gtctcagtgt tctcctaacc cccgccccc taacaaggtc     900
ccaggcagtc agctctttgt cagctctcag gcccctgagc ggccaagcag agcccacggg     960
cgtgagaacc ataccacact gcccgctgac tggaagcacc gccgggagcc ccatgacaga    1020
ggcgccagcc acctggaccg aagctatagc tgtagctata gccacaggaa tggcccagga    1080
ccattctgtc ataaaggtcc catctctgaa gagggactag gggcaagcgt tatgtccctg    1140
agcagtgaga acccctatgc taccatccga gacctgccca gcctgcctgg ggaaccccga    1200
gaaagtggct atgtggagat gaaaggacct ccatcagtgt cccctcccag gcagtctctt    1260
catctccggg acaggcagca gcggcaactg cagccacaga gggacagcgg cacctatgag    1320
cagcccagcc ccttgagcca taatgaagag tctttgggct ccacgccccc gcttcctcca    1380
ggcctgcctc ctggtcacta cgactccccc aagaacagcc atatccctgg acactatgac    1440
ttgcctccag tacggcatcc tccatccccc ccatcccggc gccaggaccg c             1491

      <210> 19
      <211> 3567
      <212> DNA
      <213> Rauttus sp.

      <220>
      <221> CDS
      <222> (925)...(2832)

      <400> 19
gtccgaccca cgcgtccgag ccacaccctg aaggtggttg gaaggaggga aggatctagg      60
tcctgagcac tggaattccc cagaacagca tctggcttcc cagacccatg ctggccacca     120
ctgatgtgtc cttccggctg ctggctgcag tgctgttctg ttgttgggtg ccctgtggca     180
ggcttgtgca atgccactct gtcccctcct cctcctggcc ctaggcctgc gtctggctgg     240
aacactcaac tccaatgatc ccaatgtctg taccttctgg gaaagcttca ccacgaccac     300
taaggagtcc caccttcgcc ccttcagcct gcccccagcc gagtcctgcg acaggccctg     360
ggaagacccc cacacctgcg ctcagcctac ggttgtctac cggactgtgt accgtcaggt     420
ggtgaagatg gactcccgcc cacgcctgca gtgctgtggg ggttactacg agagcagtgg     480
agcctgtgtc ccactctgtg cccaggagtg tgtccacggt cgctgtgtgg ctcctaatcg     540
gtgccagtgt gcaccaggct ggcggggtga cgactgttcc agtgagtgtg ctcctggaat     600
gtggggacca cagtgtgaca ggctctgcct ctgtggcaac agcagttcct gtgatcccag     660
gagtggggtg tgtttttgcc cctctggcct gcagccccca gactgccttc agccttgccc     720
cgatggccac tatggtcctg cctgccagtt tgattgccat tgctatgggg catcctgtga     780
ccccccgggat ggagcctgct tctgcccccc agggagaaca ggacccaggg cactgatggc     840
ttcttctgcc ccagaactta tccttgccaa aatggaggtg ttcctcaggg ctctcaaggc     900
tcctgcagct gcccaccggg ctgg atg ggt gtc atc tgt tcc ctg cca tgc        951
                          Met Gly Val Ile Cys Ser Leu Pro Cys
                           1               5

cca gag ggt ttc cac gga ccc aac tgt act cag gaa tgt cgt tgc cac       999
Pro Glu Gly Phe His Gly Pro Asn Cys Thr Gln Glu Cys Arg Cys His
 10                  15                  20                  25

aat ggt ggc ctt tgt gac agg ttt act ggg cag tgc cac tgt gct cct      1047
Asn Gly Gly Leu Cys Asp Arg Phe Thr Gly Gln Cys His Cys Ala Pro
                 30                  35                  40

ggc tat atc ggg gat cgg tgc cgt gaa gag tgc cct gtg ggc cgc ttc      1095
Gly Tyr Ile Gly Asp Arg Cys Arg Glu Glu Cys Pro Val Gly Arg Phe
```

```
                45                  50                  55

ggt caa gac tgt gct gag acc tgt gac tgt gct cct ggc gct cgt tgc     1143
Gly Gln Asp Cys Ala Glu Thr Cys Asp Cys Ala Pro Gly Ala Arg Cys
        60                  65                  70

ttt cct gcc aat ggc gcg tgt ctg tgc gaa cat ggc ttc aca ggc gac     1191
Phe Pro Ala Asn Gly Ala Cys Leu Cys Glu His Gly Phe Thr Gly Asp
    75                  80                  85

cgc tgc act gag cga ctc tgt cca gat ggc cgc tat ggt ctg agc tgc     1239
Arg Cys Thr Glu Arg Leu Cys Pro Asp Gly Arg Tyr Gly Leu Ser Cys
90                  95                  100                 105

caa gat ccc tgc acc tgc gac cca gaa cac agt ctc agc tgc cac cca     1287
Gln Asp Pro Cys Thr Cys Asp Pro Glu His Ser Leu Ser Cys His Pro
                110                 115                 120

atg cac ggc gag tgc tcc tgc cag cca ggt tgg gcg ggc ctc cac tgc     1335
Met His Gly Glu Cys Ser Cys Gln Pro Gly Trp Ala Gly Leu His Cys
            125                 130                 135

aac gag agc tgc cct cag gac acg cac gga gcc ggt tgc cag gag cac     1383
Asn Glu Ser Cys Pro Gln Asp Thr His Gly Ala Gly Cys Gln Glu His
        140                 145                 150

tgc ctc tgt ctg cac ggc ggt gtt tgc ctc gcc gac agc ggc ctc tgc     1431
Cys Leu Cys Leu His Gly Gly Val Cys Leu Ala Asp Ser Gly Leu Cys
    155                 160                 165

cgg tgt gca cct ggc tac acg gga cct cac tgc gct aat ctt tgt cca     1479
Arg Cys Ala Pro Gly Tyr Thr Gly Pro His Cys Ala Asn Leu Cys Pro
170                 175                 180                 185

cct aac act tat ggg atc aac tgt tcc tcc cac tgc tcc tgt gaa aat     1527
Pro Asn Thr Tyr Gly Ile Asn Cys Ser Ser His Cys Ser Cys Glu Asn
                190                 195                 200

gcc att gcc tgc tct cct gtc gac ggc acg tgc atc tgc aag gaa ggt     1575
Ala Ile Ala Cys Ser Pro Val Asp Gly Thr Cys Ile Cys Lys Glu Gly
            205                 210                 215

tgg cag cgt ggt aac tgc tct gtg ccc tgt ccc cct ggc acc tgg ggc     1623
Trp Gln Arg Gly Asn Cys Ser Val Pro Cys Pro Pro Gly Thr Trp Gly
        220                 225                 230

ttc agt tgc aat gcc agt tgc cag tgt gcc cac gag gga gtc tgc agc     1671
Phe Ser Cys Asn Ala Ser Cys Gln Cys Ala His Glu Gly Val Cys Ser
    235                 240                 245

ccc caa act gga gcc tgt act tgc acc cct ggg tgg cgt ggg gtt cac     1719
Pro Gln Thr Gly Ala Cys Thr Cys Thr Pro Gly Trp Arg Gly Val His
250                 255                 260                 265

tgc caa ctt ccg tgc ccg aag gga cag ttt ggt gaa ggt tgt gcc agt     1767
Cys Gln Leu Pro Cys Pro Lys Gly Gln Phe Gly Glu Gly Cys Ala Ser
                270                 275                 280
```

```
gtc tgt gac tgt gac cac tcc gat ggc tgt gac cct gtt cat gga cac     1815
Val Cys Asp Cys Asp His Ser Asp Gly Cys Asp Pro Val His Gly His
            285                 290                 295

tgc cga tgt cag gct ggc tgg atg ggc aca cgt tgc cac ctg cct tgc     1863
Cys Arg Cys Gln Ala Gly Trp Met Gly Thr Arg Cys His Leu Pro Cys
        300                 305                 310

cca gag ggc ttt tgg gga gcc aac tgc agc aat gcc tgt acc tgc aag     1911
Pro Glu Gly Phe Trp Gly Ala Asn Cys Ser Asn Ala Cys Thr Cys Lys
    315                 320                 325

aat ggt ggc act tgt gta cct gag aac ggc aac tgt gtg tgc gca cca     1959
Asn Gly Gly Thr Cys Val Pro Glu Asn Gly Asn Cys Val Cys Ala Pro
330                 335                 340                 345

ggg ttc aga ggc ccc tcc tgc cag agg ccc tgc ccg cct ggt cgc tat     2007
Gly Phe Arg Gly Pro Ser Cys Gln Arg Pro Cys Pro Pro Gly Arg Tyr
                350                 355                 360

ggc aaa cgc tgt gtg ccc tgc aag tgc aac aac cat tct tcc tgc cac     2055
Gly Lys Arg Cys Val Pro Cys Lys Cys Asn Asn His Ser Ser Cys His
            365                 370                 375

ccg tcg gat ggg acc tgc tcc tgc ctg gca ggc tgg aca ggc cct gac     2103
Pro Ser Asp Gly Thr Cys Ser Cys Leu Ala Gly Trp Thr Gly Pro Asp
        380                 385                 390

tgc tct gaa tca tgt ccc cca ggc cac tgg gga ctc aaa tgc tcc caa     2151
Cys Ser Glu Ser Cys Pro Pro Gly His Trp Gly Leu Lys Cys Ser Gln
    395                 400                 405

ccc tgc cag tgt cat cat ggt gcc acc tgc cac ccc cag gat ggg agc     2199
Pro Cys Gln Cys His His Gly Ala Thr Cys His Pro Gln Asp Gly Ser
410                 415                 420                 425

tgt gtc tgc atc cca ggc tgg act gga ccc aac tgc tcg gaa ggc tgc     2247
Cys Val Cys Ile Pro Gly Trp Thr Gly Pro Asn Cys Ser Glu Gly Cys
                430                 435                 440

cca tca aga atg ttt ggt gtc aac tgc tcc cag cta tgt cag tgt gat     2295
Pro Ser Arg Met Phe Gly Val Asn Cys Ser Gln Leu Cys Gln Cys Asp
            445                 450                 455

cct gga gag atg tgc cac cca gag act ggg gct tgc gtc tgt ccc cca     2343
Pro Gly Glu Met Cys His Pro Glu Thr Gly Ala Cys Val Cys Pro Pro
        460                 465                 470

gga cac agt ggt gcg cac tgc aaa gtg ggc agc cag gag tcc ttc acc     2391
Gly His Ser Gly Ala His Cys Lys Val Gly Ser Gln Glu Ser Phe Thr
    475                 480                 485

ata atg ccc acc tct cct gtg atc cat aac tca ctg ggt gcc gtg att     2439
Ile Met Pro Thr Ser Pro Val Ile His Asn Ser Leu Gly Ala Val Ile
490                 495                 500                 505

ggc att gca gtg ctg ggg acc ctt gtg gtg gcc ctg gta gca ctg ttt     2487
Gly Ile Ala Val Leu Gly Thr Leu Val Val Ala Leu Val Ala Leu Phe
```

```
                510                 515                 520

att ggc tac cga cac tgg caa aag ggc aag gaa cat gag cac ttg gca     2535
Ile Gly Tyr Arg His Trp Gln Lys Gly Lys Glu His Glu His Leu Ala
            525                 530                 535

gtg gct tac agc act ggg cga ctg gat ggc tcc gat tac gtc atg cca     2583
Val Ala Tyr Ser Thr Gly Arg Leu Asp Gly Ser Asp Tyr Val Met Pro
        540                 545                 550

gat gtc tct ccg agc tac agt cac tac tat tcc aac cct agc tac cac     2631
Asp Val Ser Pro Ser Tyr Ser His Tyr Tyr Ser Asn Pro Ser Tyr His
    555                 560                 565

aca ctg tct cag tgt tct cct aac cct cca ccc cct aac aag att cca     2679
Thr Leu Ser Gln Cys Ser Pro Asn Pro Pro Pro Pro Asn Lys Ile Pro
570                 575                 580                 585

ggc agt cag ctg ttt gtc agc tcc cag gca tct gag cgg cca aac aga     2727
Gly Ser Gln Leu Phe Val Ser Ser Gln Ala Ser Glu Arg Pro Asn Arg
                590                 595                 600

aac cat ggg cga gat aac cac gcc aca ctg ccc gct gac tgg aag cac     2775
Asn His Gly Arg Asp Asn His Ala Thr Leu Pro Ala Asp Trp Lys His
            605                 610                 615

cga cgg gag tcc cat gac aga gct ttc ctc agg cac cag cca cct gga     2823
Arg Arg Glu Ser His Asp Arg Ala Phe Leu Arg His Gln Pro Pro Gly
        620                 625                 630

ccg aag gta tagcttgtagc tatggccaca ggaatggccc ggggccattc             2872
Pro Lys Val
    635

tgtcataaag gtcccatctc tgaagaagga ctaggggcaa gcgttatgtc cctgagcagt   2932
gagaacccct atgcgaccat ccgagacctg cccggcctgc ctggggaacc ccgagaaagc   2992
agctatgtgg agatgaaagg ccctccatca gtgtctcccc ccaggcagcc tcttcatctc   3052
cgggacaggc agcagcagca actgcagtct cagagagaca gcggcaccta tgagcagccc   3112
actcccttga gccgtaatga agagtctgtg ggctccatgc cccctcttcc tccgggcctg   3172
ccacccggcc actatgactc gcccaaaaac agccacatcc ctggacacta tgacttgcct   3232
ccagtacggc atcctccatc acctccatcc cggcgccagg accgctgagg agccagcatg   3292
gtatgggaga gtgcctgtga accctgccag gagcagggcc tggaccagca ggccatgaat   3352
agacatactt ggtgaagtga acggagactg aggatggctc tgcttccacc gagggagaca   3412
ctagttggca aagtgtctaa cctccctttt ccagcccatt gctcaagtcc cccaggctgt   3472
ggacatgagc tggtgggcag aatgttgttg ttgaagtctg attttagatt gatttttaa    3532
aaaaaaaaaa aaaaaaaaaa aaaaagggcg gccgc                              3567

      <210> 20
      <211> 636
      <212> PRT
      <213> Rauttus sp.

      <400> 20
Met Gly Val Ile Cys Ser Leu Pro Cys Pro Glu Gly Phe His Gly Pro
 1               5                  10                  15
Asn Cys Thr Gln Glu Cys Arg Cys His Asn Gly Gly Leu Cys Asp Arg
            20                  25                  30
Phe Thr Gly Gln Cys His Cys Ala Pro Gly Tyr Ile Gly Asp Arg Cys
```

```
          35                  40                  45
Arg Glu Glu Cys Pro Val Gly Arg Phe Gly Gln Asp Cys Ala Glu Thr
    50                  55                  60
Cys Asp Cys Ala Pro Gly Ala Arg Cys Phe Pro Ala Asn Gly Ala Cys
65                  70                  75                  80
Leu Cys Glu His Gly Phe Thr Gly Asp Arg Cys Thr Glu Arg Leu Cys
                85                  90                  95
Pro Asp Gly Arg Tyr Gly Leu Ser Cys Gln Asp Pro Cys Thr Cys Asp
            100                 105                 110
Pro Glu His Ser Leu Ser Cys His Pro Met His Gly Glu Cys Ser Cys
        115                 120                 125
Gln Pro Gly Trp Ala Gly Leu His Cys Asn Glu Ser Cys Pro Gln Asp
    130                 135                 140
Thr His Gly Ala Gly Cys Gln Glu His Cys Leu Cys Leu His Gly Gly
145                 150                 155                 160
Val Cys Leu Ala Asp Ser Gly Leu Cys Arg Cys Ala Pro Gly Tyr Thr
                165                 170                 175
Gly Pro His Cys Ala Asn Leu Cys Pro Pro Asn Thr Tyr Gly Ile Asn
            180                 185                 190
Cys Ser Ser His Cys Ser Cys Glu Asn Ala Ile Ala Cys Ser Pro Val
        195                 200                 205
Asp Gly Thr Cys Ile Cys Lys Glu Gly Trp Gln Arg Gly Asn Cys Ser
    210                 215                 220
Val Pro Cys Pro Pro Gly Thr Trp Gly Phe Ser Cys Asn Ala Ser Cys
225                 230                 235                 240
Gln Cys Ala His Glu Gly Val Cys Ser Pro Gln Thr Gly Ala Cys Thr
                245                 250                 255
Cys Thr Pro Gly Trp Arg Gly Val His Cys Gln Leu Pro Cys Pro Lys
            260                 265                 270
Gly Gln Phe Gly Glu Gly Cys Ala Ser Val Cys Asp Cys Asp His Ser
        275                 280                 285
Asp Gly Cys Asp Pro Val His Gly His Cys Arg Cys Gln Ala Gly Trp
    290                 295                 300
Met Gly Thr Arg Cys His Leu Pro Cys Pro Glu Gly Phe Trp Gly Ala
305                 310                 315                 320
Asn Cys Ser Asn Ala Cys Thr Cys Lys Asn Gly Gly Thr Cys Val Pro
                325                 330                 335
Glu Asn Gly Asn Cys Val Cys Ala Pro Gly Phe Arg Gly Pro Ser Cys
            340                 345                 350
Gln Arg Pro Cys Pro Pro Gly Arg Tyr Gly Lys Arg Cys Val Pro Cys
        355                 360                 365
Lys Cys Asn Asn His Ser Ser Cys His Pro Ser Asp Gly Thr Cys Ser
    370                 375                 380
Cys Leu Ala Gly Trp Thr Gly Pro Asp Cys Ser Glu Ser Cys Pro Pro
385                 390                 395                 400
Gly His Trp Gly Leu Lys Cys Ser Gln Pro Cys Gln Cys His His Gly
                405                 410                 415
Ala Thr Cys His Pro Gln Asp Gly Ser Cys Val Cys Ile Pro Gly Trp
            420                 425                 430
Thr Gly Pro Asn Cys Ser Glu Gly Cys Pro Ser Arg Met Phe Gly Val
        435                 440                 445
Asn Cys Ser Gln Leu Cys Gln Cys Asp Pro Gly Glu Met Cys His Pro
    450                 455                 460
Glu Thr Gly Ala Cys Val Cys Pro Pro Gly His Ser Gly Ala His Cys
465                 470                 475                 480
Lys Val Gly Ser Gln Glu Ser Phe Thr Ile Met Pro Thr Ser Pro Val
                485                 490                 495
Ile His Asn Ser Leu Gly Ala Val Ile Gly Ile Ala Val Leu Gly Thr
```

```
            500                 505                 510
Leu Val Val Ala Leu Val Ala Leu Phe Ile Gly Tyr Arg His Trp Gln
        515                 520                 525
Lys Gly Lys Glu His Glu His Leu Ala Val Ala Tyr Ser Thr Gly Arg
    530                 535                 540
Leu Asp Gly Ser Asp Tyr Val Met Pro Asp Val Ser Pro Ser Tyr Ser
545                 550                 555                 560
His Tyr Tyr Ser Asn Pro Ser Tyr His Thr Leu Ser Gln Cys Ser Pro
                565                 570                 575
Asn Pro Pro Pro Pro Asn Lys Ile Pro Gly Ser Gln Leu Phe Val Ser
            580                 585                 590
Ser Gln Ala Ser Glu Arg Pro Asn Arg Asn His Gly Arg Asp Asn His
        595                 600                 605
Ala Thr Leu Pro Ala Asp Trp Lys His Arg Arg Glu Ser His Asp Arg
    610                 615                 620
Ala Phe Leu Arg His Gln Pro Pro Gly Pro Lys Val
625                 630                 635
```

<210> 21
<211> 1908
<212> DNA
<213> Rattus sp.

<400> 21

```
atgggtgtca tctgttccct gccatgccca gagggtttcc acggacccaa ctgtactcag      60
gaatgtcgtt gccacaatgg tggcctttgt gacaggttta ctgggcagtg ccactgtgct     120
cctggctata tcggggatcg gtgccgtgaa gagtgccctg tgggccgctt cggtcaagac     180
tgtgctgaga cctgtgactg tgctcctggc gctcgttgct ttcctgccaa tggcgcgtgt     240
ctgtgcgaac atggcttcac aggcgaccgc tgcactgagc gactctgtcc agatggccgc     300
tatggtctga gctgccaaga tccctgcacc tgcgacccag aacacagtct cagctgccac     360
ccaatgcacg gcgagtgctc ctgccagcca ggttgggcgg gcctccactg caacgagagc     420
tgccctcagg acacgcacgg agccggttgc caggagcact gcctctgtct gcacggcggt     480
gtttgcctcg ccgacagcgg cctctgccgg tgtgcacctg gctacacggg acctcactgc     540
gctaatcttt gtccacctaa cacttatggg atcaactgtt cctcccactg ctcctgtgaa     600
aatgccattg cctgctctcc tgtcgacggc acgtgcatct gcaaggaagg ttggcagcgt     660
ggtaactgct ctgtgccctg tcccctggc acctggggct tcagttgcaa tgccagttgc     720
cagtgtgccc acgagggagt ctgcagcccc caaactggag cctgtacttg caccctggg      780
tggcgtgggg ttcactgcca acttccgtgc ccgaagggac agtttggtga aggttgtgcc     840
agtgtctgtg actgtgacca ctccgatggc tgtgaccctg ttcatggaca ctgccgatgt     900
caggctggct ggatgggcac acgttgccac ctgccttgcc cagagggctt ttggggagcc     960
aactgcagca atgcctgtac ctgcaagaat ggtggcactt gtgtacctga gaacggcaac    1020
tgtgtgtgcg caccagggtt cagaggcccc tcctgccaga ggccctgccc gcctggtcgc    1080
tatggcaaac gctgtgtgcc ctgcaagtgc aacaaccatt cttcctgcca cccgtcggat    1140
gggacctgct cctgcctggc aggctggaca ggccctgact gctctgaatc atgtccccca    1200
ggccactggg gactcaaatg ctcccaaccc tgccagtgtc atcatggtgc cacctgccac    1260
ccccaggatg ggagctgtgt ctgcatccca ggctggactg gacccaactg ctcggaaggc    1320
tgcccatcaa gaatgtttgg tgtcaactgc tcccagctat gtcagtgtga tcctggagag    1380
atgtgccacc cagagactgg ggcttgcgtc tgtccccag gacacagtgg tgcgcactgc     1440
aaagtgggca gccaggagtc cttcaccata atgcccacct ctcctgtgat ccataactca    1500
ctgggtgccg tgattggcat tgcagtgctg gggacccttg tggtggccct ggtagcactg    1560
tttattggct accgacactg gcaaaagggc aaggaacatg agcacttggc agtggcttac    1620
agcactgggc gactggatgg ctccgattac gtcatgccag atgtctctcc gagctacagt    1680
cactactatt ccaaccctag ctaccacaca ctgtctcagt gttctcctaa ccctccaccc    1740
cctaacaaga ttccaggcag tcagctgttt gtcagctccc aggcatctga gcggccaaac    1800
agaaaccatg ggcgagataa ccacgccaca ctgcccgctg actggaagca ccgacgggag    1860
tcccatgaca gagctttcct caggcaccag ccacctggac cgaaggta                1908
```

<210> 22

<211> 1497
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (217)...(684)

<400> 22

```
gtcgacccac gcgtccggct cccagcccac ccccaaacag acacagcgta gcccgggcca     60
gctcttaagg agttcaggag tgagaagagg ccctcagaga tctgacagcc taggagtgcg    120
tggacaccac ctcagcccac tgagcaggag tcacagcacg aagaccaagc gcaaagcgac    180
ccctgccctc catcctgact gctcctccta agagag atg gca ccg gcc aga gca      234
                                          Met Ala Pro Ala Arg Ala
                                           1               5

gga ttc tgc ccc ctt ctg ctg ctt ctg ctg ctg ggg ctg tgg gtg gca      282
Gly Phe Cys Pro Leu Leu Leu Leu Leu Leu Leu Gly Leu Trp Val Ala
            10                  15                  20

gag atc cca gtc agt gcc aag ccc aag ggc atg acc tca tca cag tgg      330
Glu Ile Pro Val Ser Ala Lys Pro Lys Gly Met Thr Ser Ser Gln Trp
         25                  30                  35

ttt aaa att cag cac atg cag ccc agc cct caa gca tgc aac tca gcc      378
Phe Lys Ile Gln His Met Gln Pro Ser Pro Gln Ala Cys Asn Ser Ala
     40                  45                  50

atg aaa aac att aac aag cac aca aaa cgg tgc aaa gac ctc aac acc      426
Met Lys Asn Ile Asn Lys His Thr Lys Arg Cys Lys Asp Leu Asn Thr
 55                  60                  65                  70

ttc ctg cac gag cct ttc tcc agt gtg gcc gcc acc tgc cag acc ccc      474
Phe Leu His Glu Pro Phe Ser Ser Val Ala Ala Thr Cys Gln Thr Pro
                75                  80                  85

aaa ata gcc tgc aag aat ggc gat aaa aac tgc cac cag agc cac ggg      522
Lys Ile Ala Cys Lys Asn Gly Asp Lys Asn Cys His Gln Ser His Gly
            90                  95                  100

ccc gtg tcc ctg acc atg tgt aag ctc acc tca ggg aag tat ccg aac      570
Pro Val Ser Leu Thr Met Cys Lys Leu Thr Ser Gly Lys Tyr Pro Asn
        105                 110                 115

tgc agg tac aaa gag aag cga cag aac aag tct tac gta gtg gcc tgt      618
Cys Arg Tyr Lys Glu Lys Arg Gln Asn Lys Ser Tyr Val Val Ala Cys
    120                 125                 130

aag cct ccc cag aaa aag gac tct cag caa ttc cac ctg gtt cct gta      666
Lys Pro Pro Gln Lys Lys Asp Ser Gln Gln Phe His Leu Val Pro Val
135                 140                 145                 150

cac ttg gac aga gtc ctt taggtttcca gactggcttg ctctttggct             714
His Leu Asp Arg Val Leu
                155

gaccttcaat tccctctcca ggactccgca ccactccccc acacccagag cattctcttc    774
ccctcatctc ttggggctgt tcctggttca gcctctgctg ggaggctgaa gctgacactc    834
```

```
tggtgagctg agctctagag ggatggcttt tcatcttttt gttgctgttt tcccagatgc    894
ttatccccaa gaaacagcaa gctcaggtct gtgggttccc tggtctatgc cattgcacat    954
gtctcccctg ccccctggca ttagggcagc atgacaagga gaggaaataa atggaaaggg   1014
ggcatatggg atttgtggac acagctgttt ctgttcctga actagaagtc ttccccagct   1074
ctgacgtggc agtgaggtga cctgaaggaa agaaaaatat aaataaatac cacttcatat   1134
ttgtatagaa tcctctaatc ccttgtgaca tagacttgac agggattgta tgccttcttt   1194
atggatgagg aaattaaggt tttagaaagc ttaatgaatt aaagagcttg tctaattagt   1254
tagtagcaga acctggactt gaacctaggt ctccttgctc taaatacagt gtaccttcta   1314
ctctaccagt tgcgcaagaa agaagtcact gttacagagg caagcggtga actaggtaag   1374
agttcactca tgaagaaacg agtgctctga agagccagtt accctgtgtt ggctgcaata   1434
aaggtcatta cctctctagc caaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa   1494
aaa                                                                 1497
```

<210> 23
<211> 156
<212> PRT
<213> Homo sapiens

<400> 23

```
Met Ala Pro Ala Arg Ala Gly Phe Cys Pro Leu Leu Leu Leu Leu Leu
 1               5                  10                  15
Leu Gly Leu Trp Val Ala Glu Ile Pro Val Ser Ala Lys Pro Lys Gly
            20                  25                  30
Met Thr Ser Ser Gln Trp Phe Lys Ile Gln His Met Gln Pro Ser Pro
        35                  40                  45
Gln Ala Cys Asn Ser Ala Met Lys Asn Ile Asn Lys His Thr Lys Arg
    50                  55                  60
Cys Lys Asp Leu Asn Thr Phe Leu His Glu Pro Phe Ser Ser Val Ala
65                  70                  75                  80
Ala Thr Cys Gln Thr Pro Lys Ile Ala Cys Lys Asn Gly Asp Lys Asn
                85                  90                  95
Cys His Gln Ser His Gly Pro Val Ser Leu Thr Met Cys Lys Leu Thr
            100                 105                 110
Ser Gly Lys Tyr Pro Asn Cys Arg Tyr Lys Glu Lys Arg Gln Asn Lys
        115                 120                 125
Ser Tyr Val Val Ala Cys Lys Pro Pro Gln Lys Lys Asp Ser Gln Gln
    130                 135                 140
Phe His Leu Val Pro Val His Leu Asp Arg Val Leu
145                 150                 155
```

<210> 24
<211> 468
<212> DNA
<213> Homo sapiens

<400> 24

```
atggcaccgg ccagagcagg attctgcccc cttctgctgc ttctgctgct ggggctgtgg      60
gtggcagaga tcccagtcag tgccaagccc aagggcatga cctcatcaca gtggtttaaa     120
attcagcaca tgcagcccag ccctcaagca tgcaactcag ccatgaaaaa cattaacaag     180
cacacaaaac ggtgcaaaga cctcaacacc ttcctgcacg agcctttctc cagtgtggcc     240
gccacctgcc agaccccaa aatagcctgc aagaatggcg ataaaaactg ccaccagagc      300
cacgggcccg tgtccctgac catgtgtaag ctcacctcag ggaagtatcc gaactgcagg     360
tacaaagaga agcgacagaa caagtcttac gtagtggcct gtaagcctcc ccagaaaaag     420
gactctcagc aattccacct ggttcctgta cacttggaca gagtcctt                  468
```

<210> 25
<211> 1788
<212> DNA

<213> Homo sapiens

<220>
<221> CDS
<222> (62)...(976)

<400> 25

```
gtcgacccac ggcgtccggc caggctccac tgaggggaac ggggacctgt ctgaagagaa      60
g atg ccc ctg ctg aca ctc tac ctg ctc ctc ttc tgg ctc tca ggc tac     109
  Met Pro Leu Leu Thr Leu Tyr Leu Leu Leu Phe Trp Leu Ser Gly Tyr
   1               5                  10                  15

tcc att gcc act caa atc acc ggt cca aca aca gtg aat ggc ttg gag       157
Ser Ile Ala Thr Gln Ile Thr Gly Pro Thr Thr Val Asn Gly Leu Glu
            20                  25                  30

cgg ggc tcc ttg acc gtg cag tgt gtt tac aga tca ggc tgg gag acc       205
Arg Gly Ser Leu Thr Val Gln Cys Val Tyr Arg Ser Gly Trp Glu Thr
         35                  40                  45

tac ttg aag tgg tgg tgt cga gga gct att tgg cgt gac tgc aag atc       253
Tyr Leu Lys Trp Trp Cys Arg Gly Ala Ile Trp Arg Asp Cys Lys Ile
     50                  55                  60

ctt gtt aaa acc agt ggg tca gag cag gag gtg aag agg gac cgg gtg       301
Leu Val Lys Thr Ser Gly Ser Glu Gln Glu Val Lys Arg Asp Arg Val
 65                  70                  75                  80

tcc atc aag gac aat cag aaa aac cgc acg ttc act gtg acc atg gag       349
Ser Ile Lys Asp Asn Gln Lys Asn Arg Thr Phe Thr Val Thr Met Glu
                85                  90                  95

gat ctc atg aaa act gat gct gac act tac tgg tgt gga att gag aaa       397
Asp Leu Met Lys Thr Asp Ala Asp Thr Tyr Trp Cys Gly Ile Glu Lys
            100                 105                 110

act gga aat gac ctt ggg gtc aca gtt caa gtg acc att gac cca gcg       445
Thr Gly Asn Asp Leu Gly Val Thr Val Gln Val Thr Ile Asp Pro Ala
        115                 120                 125

tcg act cct gcc ccc acc acg cct act tcc act acg ttt aca gca cca       493
Ser Thr Pro Ala Pro Thr Thr Pro Thr Ser Thr Thr Phe Thr Ala Pro
    130                 135                 140

gtc acc caa gaa gaa act agc agc tcc cca act ctg acc ggc cac cac       541
Val Thr Gln Glu Glu Thr Ser Ser Ser Pro Thr Leu Thr Gly His His
145                 150                 155                 160

ttg gac aac agg cac aag ctc ctg aag ctc agt gtc ctc ctg ccc ctc       589
Leu Asp Asn Arg His Lys Leu Leu Lys Leu Ser Val Leu Leu Pro Leu
                165                 170                 175

atc ttc acc ata ttg ctg ctg ctt ttg gtg gcc gcc tca ctc ttg gct       637
Ile Phe Thr Ile Leu Leu Leu Leu Leu Val Ala Ala Ser Leu Leu Ala
            180                 185                 190

tgg agg atg atg aag tac cag cag aaa gca gcc ggg atg tcc cca gag       685
Trp Arg Met Met Lys Tyr Gln Gln Lys Ala Ala Gly Met Ser Pro Glu
```

```
        195                 200                 205

cag gta ctg cag ccc ctg gag ggc gac ctc tgc tat gca gac ctg acc      733
Gln Val Leu Gln Pro Leu Glu Gly Asp Leu Cys Tyr Ala Asp Leu Thr
    210                 215                 220

ctg cag ctg gcc gga acc tcc ccg cga aag gct acc acg aag ctt tcc      781
Leu Gln Leu Ala Gly Thr Ser Pro Arg Lys Ala Thr Thr Lys Leu Ser
225                 230                 235                 240

tct gcc cag gtt gac cag gtg gaa gtg gaa tat gtc acc atg gct tcc      829
Ser Ala Gln Val Asp Gln Val Glu Val Glu Tyr Val Thr Met Ala Ser
                245                 250                 255

ttg ccg aag gag gac att tcc tat gca tct ctg acc ttg ggt gct gag      877
Leu Pro Lys Glu Asp Ile Ser Tyr Ala Ser Leu Thr Leu Gly Ala Glu
            260                 265                 270

gat cag gaa ccg acc tac tgc aac atg ggc cac ctc agt agc cac ctc      925
Asp Gln Glu Pro Thr Tyr Cys Asn Met Gly His Leu Ser Ser His Leu
        275                 280                 285

ccc ggc agg ggc cct gag gag ccc acg gaa tac agc acc atc agc agg      973
Pro Gly Arg Gly Pro Glu Glu Pro Thr Glu Tyr Ser Thr Ile Ser Arg
    290                 295                 300

cct tagcctgcac tccaggctcc ttcttggacc ccaggctgtg agcacactcc          1026
Pro
305

tgcctcatcg accgtctgcc ccctgctccc ctcatcagga ccaacccggg gactggtgcc   1086
tctgcctgat cagccagcat tgcccctagc tctgggttgg gcttggggcc aagtctcagg   1146
gggcttctag gagttggggt tttctaaacg tcccctcctc tcctacatag ttgaggaggg   1206
ggctagggat atgctctggg gctttcatgg gaatgatgaa gatgataatg agaaaaatgt   1266
tatcattatt atcatgaagt accattatca taatacaatg aacctttatt tattgcctac   1326
cacatgttat gggctgaata atggcccccа aagatatctg tgtcctaatc ctcagaactt   1386
gtgactgtta ccttctgtgg cagaaaggga cagtgcagat gtatgtaagt taaggacttt   1446
gagatagaga ggttattctt gctgattcag gtgggcccaa aatatcacca caagggtcct   1506
cataagaaag aggccagaag gtcaaagagg tagagacaaa gtgatgatgg aagtggacgt   1566
gggtgtgacg tgagcagggg ccatgaatgc cgcagccttc agatgccaga aagggaaagg   1626
aatggattcc cctgcctgga gcctccaaaa gaaaccagcc ctgcccacgc cttgacttga   1686
gcccattgaa actgatcttg agctcctggc ctccagaatt gcaggagaat aaatttgtgt   1746
tgtttttaaa aaaaaaaaaa aaaaaaaagg gcggccgcta ga                      1788
```

<210> 26
<211> 305
<212> PRT
<213> Homo sapiens

<400> 26

```
Met Pro Leu Leu Thr Leu Tyr Leu Leu Leu Phe Trp Leu Ser Gly Tyr
 1               5                  10                  15
Ser Ile Ala Thr Gln Ile Thr Gly Pro Thr Thr Val Asn Gly Leu Glu
            20                  25                  30
Arg Gly Ser Leu Thr Val Gln Cys Val Tyr Arg Ser Gly Trp Glu Thr
        35                  40                  45
Tyr Leu Lys Trp Trp Cys Arg Gly Ala Ile Trp Arg Asp Cys Lys Ile
    50                  55                  60
```

```
Leu Val Lys Thr Ser Gly Ser Glu Gln Glu Val Lys Arg Asp Arg Val
65                  70                  75                  80
Ser Ile Lys Asp Asn Gln Lys Asn Arg Thr Phe Thr Val Thr Met Glu
                85                  90                  95
Asp Leu Met Lys Thr Asp Ala Asp Thr Tyr Trp Cys Gly Ile Glu Lys
            100                 105                 110
Thr Gly Asn Asp Leu Gly Val Thr Val Gln Val Thr Ile Asp Pro Ala
        115                 120                 125
Ser Thr Pro Ala Pro Thr Thr Pro Thr Ser Thr Thr Phe Thr Ala Pro
    130                 135                 140
Val Thr Gln Glu Glu Thr Ser Ser Ser Pro Thr Leu Thr Gly His His
145                 150                 155                 160
Leu Asp Asn Arg His Lys Leu Leu Lys Leu Ser Val Leu Leu Pro Leu
                165                 170                 175
Ile Phe Thr Ile Leu Leu Leu Leu Leu Val Ala Ala Ser Leu Leu Ala
            180                 185                 190
Trp Arg Met Met Lys Tyr Gln Gln Lys Ala Ala Gly Met Ser Pro Glu
        195                 200                 205
Gln Val Leu Gln Pro Leu Glu Gly Asp Leu Cys Tyr Ala Asp Leu Thr
    210                 215                 220
Leu Gln Leu Ala Gly Thr Ser Pro Arg Lys Ala Thr Thr Lys Leu Ser
225                 230                 235                 240
Ser Ala Gln Val Asp Gln Val Glu Val Glu Tyr Val Thr Met Ala Ser
                245                 250                 255
Leu Pro Lys Glu Asp Ile Ser Tyr Ala Ser Leu Thr Leu Gly Ala Glu
            260                 265                 270
Asp Gln Glu Pro Thr Tyr Cys Asn Met Gly His Leu Ser Ser His Leu
        275                 280                 285
Pro Gly Arg Gly Pro Glu Glu Pro Thr Glu Tyr Ser Thr Ile Ser Arg
    290                 295                 300
Pro
305
```

<210> 27
<211> 915
<212> DNA
<213> Homo sapiens

<400> 27

```
atgcccctgc tgacactcta cctgctcctc ttctggctct caggctactc cattgccact      60
caaatcaccg gtccaacaac agtgaatggc ttggagcggg gctccttgac cgtgcagtgt     120
gtttacagat caggctggga gacctacttg aagtggtggt gtcgaggagc tatttggcgt     180
gactgcaaga tccttgttaa aaccagtggg tcagagcagg aggtgaagag ggaccgggtg     240
tccatcaagg acaatcagaa aaaccgcacg ttcactgtga ccatggagga tctcatgaaa     300
actgatgctg acacttactg gtgtggaatt gagaaaactg gaaatgacct tggggtcaca     360
gttcaagtga ccattgaccc agcgtcgact cctgccccca ccacgcctac ttccactacg     420
tttacagcac cagtcaccca agaagaaact agcagctccc caactctgac cggccaccac     480
ttggacaaca ggcacaagct cctgaagctc agtgtcctcc tgcccctcat cttcaccata     540
ttgctgctgc ttttggtggc cgcctcactc ttggcttgga ggatgatgaa gtaccagcag     600
aaagcagccg ggatgtcccc agagcaggta ctgcagcccc tggagggcga cctctgctat     660
gcagacctga ccctgcagct ggccggaacc tccccgcgaa aggctaccac gaagctttcc     720
tctgcccagg ttgaccaggt ggaagtggaa tatgtcacca tggcttcctt gccgaaggag     780
gacatttcct atgcatctct gaccttgggt gctgaggatc aggaaccgac ctactgcaac     840
atgggccacc tcagtagcca cctccccggc aggggccctg aggagcccac ggaatacagc     900
accatcagca ggcct                                                      915
```

<210> 28
<211> 3258

<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (42)...(1625)

<400> 28

```
cacgcgtccg gccagttctt ggaggagact ctgcacaggg c atg gat cac tgt ggt      56
                                                 Met Asp His Cys Gly
                                                  1               5

gcc ctt ttc ctg tgc ctg tgc ctt ctg act ttg cag aat gca aca aca      104
Ala Leu Phe Leu Cys Leu Cys Leu Leu Thr Leu Gln Asn Ala Thr Thr
                10                  15                  20

gag aca tgg gaa gaa ctc ctg agc tac atg gag aat atg cag gtg tcc      152
Glu Thr Trp Glu Glu Leu Leu Ser Tyr Met Glu Asn Met Gln Val Ser
            25                  30                  35

agg ggc cgg agc tca gtt ttt tcc tct cgt caa ctc cac cag ctg gag      200
Arg Gly Arg Ser Ser Val Phe Ser Ser Arg Gln Leu His Gln Leu Glu
        40                  45                  50

cag atg cta ctg aac acc agc ttc cca ggc tac aac ctg acc ttg cag      248
Gln Met Leu Leu Asn Thr Ser Phe Pro Gly Tyr Asn Leu Thr Leu Gln
    55                  60                  65

aca ccc acc atc cag tct ctg gcc ttc aag ctg agc tgt gac ttc tct      296
Thr Pro Thr Ile Gln Ser Leu Ala Phe Lys Leu Ser Cys Asp Phe Ser
70                  75                  80                  85

ggc ctc tcg ctg acc agt gcc act ctg aag cgg gtg ccc cag gca gga      344
Gly Leu Ser Leu Thr Ser Ala Thr Leu Lys Arg Val Pro Gln Ala Gly
                90                  95                  100

ggt cag cat gcc cgg ggt cag cac gcc atg cag ttc ccc gcc gag ctg      392
Gly Gln His Ala Arg Gly Gln His Ala Met Gln Phe Pro Ala Glu Leu
            105                 110                 115

acc cgg gac gcc tgc aag acc cgc ccc agg gag ctg cgg ctc atc tgt      440
Thr Arg Asp Ala Cys Lys Thr Arg Pro Arg Glu Leu Arg Leu Ile Cys
        120                 125                 130

atc tac ttc tcc aac acc cac ttt ttc aag gat gaa aac aac tca tct      488
Ile Tyr Phe Ser Asn Thr His Phe Phe Lys Asp Glu Asn Asn Ser Ser
    135                 140                 145

ctg ctg aat aac tac gtc ctg ggg gcc cag ctg agt cat ggg cac gtg      536
Leu Leu Asn Asn Tyr Val Leu Gly Ala Gln Leu Ser His Gly His Val
150                 155                 160                 165

aac aac ctc agg gat cct gtg aac atc agc ttc tgg cac aac caa agc      584
Asn Asn Leu Arg Asp Pro Val Asn Ile Ser Phe Trp His Asn Gln Ser
                170                 175                 180

ctg gaa ggc tac acc ctg acc tgt gtc ttc tgg aag gag gga gcc agg      632
Leu Glu Gly Tyr Thr Leu Thr Cys Val Phe Trp Lys Glu Gly Ala Arg
```

```
                185                 190                 195

aaa cag ccc tgg ggg ggc tgg agc cct gag ggc tgt cgt aca gag cag      680
Lys Gln Pro Trp Gly Gly Trp Ser Pro Glu Gly Cys Arg Thr Glu Gln
        200                 205                 210

ccc tcc cac tct cag gtg ctc tgc cgc tgc aac cac ctc acc tac ttt      728
Pro Ser His Ser Gln Val Leu Cys Arg Cys Asn His Leu Thr Tyr Phe
    215                 220                 225

gct gtt ctc atg caa ctc tcc cca gcc ctg gtc cct gca gag ttg ctg      776
Ala Val Leu Met Gln Leu Ser Pro Ala Leu Val Pro Ala Glu Leu Leu
230                 235                 240                 245

gca cct ctt acg tac atc tcc ctc gtg ggc tgc agc atc tcc atc gtg      824
Ala Pro Leu Thr Tyr Ile Ser Leu Val Gly Cys Ser Ile Ser Ile Val
                250                 255                 260

gcc tcg ctg atc aca gtc ctg ctg cac ttc cat ttc agg aag cag agt      872
Ala Ser Leu Ile Thr Val Leu Leu His Phe His Phe Arg Lys Gln Ser
            265                 270                 275

gac tcc tta aca cgc atc cac atg aac ctg cat gcc tcc gtg ctg ctc      920
Asp Ser Leu Thr Arg Ile His Met Asn Leu His Ala Ser Val Leu Leu
        280                 285                 290

ctg aac atc gcc ttc ctg ctg agc ccc gca ttc gca atg tct cct gtg      968
Leu Asn Ile Ala Phe Leu Leu Ser Pro Ala Phe Ala Met Ser Pro Val
    295                 300                 305

ccc ggg tca gca tgc acg gct ctg gcc gct gcc ctg cac tac gcg ctg     1016
Pro Gly Ser Ala Cys Thr Ala Leu Ala Ala Ala Leu His Tyr Ala Leu
310                 315                 320                 325

ctc agc tgc ctc acc tgg atg gcc atc gag ggc ttc aac ctc tac ctc     1064
Leu Ser Cys Leu Thr Trp Met Ala Ile Glu Gly Phe Asn Leu Tyr Leu
                330                 335                 340

ctc ctc ggg cgt gtc tac aac atc tac atc cgc aga tat gtg ttc aag     1112
Leu Leu Gly Arg Val Tyr Asn Ile Tyr Ile Arg Arg Tyr Val Phe Lys
            345                 350                 355

ctt ggt gtg cta ggc tgg ggg gcc cca gcc ctc ctg gtg ctg ctt tcc     1160
Leu Gly Val Leu Gly Trp Gly Ala Pro Ala Leu Leu Val Leu Leu Ser
        360                 365                 370

ctc tct gtc aag agc tcg gta tac gga ccc tgc aca atc ccc gtc ttc     1208
Leu Ser Val Lys Ser Ser Val Tyr Gly Pro Cys Thr Ile Pro Val Phe
    375                 380                 385

gac agc tgg gag aat ggc aca ggc ttc cag aac atg tcc ata tgc tgg     1256
Asp Ser Trp Glu Asn Gly Thr Gly Phe Gln Asn Met Ser Ile Cys Trp
390                 395                 400                 405

gtg cgg agc ccc gtg gtg cac agt gtc ctg gtc atg ggc tac ggc ggc     1304
Val Arg Ser Pro Val Val His Ser Val Leu Val Met Gly Tyr Gly Gly
                410                 415                 420
```

```
ctc acg tcc ctc ttc aac ctg gtg gtg ctg gcc tgg gcg ctg tgg acc     1352
Leu Thr Ser Leu Phe Asn Leu Val Val Leu Ala Trp Ala Leu Trp Thr
            425                 430                 435

ctg cgc agg ctg cgg gag cgg gcg gat gca cca agt gtc agg gcc tgc     1400
Leu Arg Arg Leu Arg Glu Arg Ala Asp Ala Pro Ser Val Arg Ala Cys
        440                 445                 450

cat gac act gtc act gtg ctg ggc ctc acc gtg ctg ctg gga acc acc     1448
His Asp Thr Val Thr Val Leu Gly Leu Thr Val Leu Leu Gly Thr Thr
    455                 460                 465

tgg gcc ttg gcc ttc ttt tct ttt ggc gtc ttc ctg ctg ccc cag ctg     1496
Trp Ala Leu Ala Phe Phe Ser Phe Gly Val Phe Leu Leu Pro Gln Leu
470                 475                 480                 485

ttc ctc ttc acc atc tta aac tcg ctc tac ggt ttc ttc ctt ttc ctg     1544
Phe Leu Phe Thr Ile Leu Asn Ser Leu Tyr Gly Phe Phe Leu Phe Leu
                490                 495                 500

tgg ttc tgc tcc cag cgg tgc cgc tca gaa gca gag gcc aag gca cag     1592
Trp Phe Cys Ser Gln Arg Cys Arg Ser Glu Ala Glu Ala Lys Ala Gln
            505                 510                 515

ata gag gcc ttc agc tcc tcc caa aca aca cag tagtccgggc ctcctggcct   1645
Ile Glu Ala Phe Ser Ser Ser Gln Thr Thr Gln
        520                 525

ggaatcctca gcctctctgg ccgccagtag cctgaggcta cggctcctgc tagagagggt   1705
ggcaggcctg ctgctggacc ccagaggcca ctgtgaccgc caaggggcct tttccacttc   1765
cacggcctct ccaggcactg aggggaaggc attgctctac ctctccctga cattttgctc   1825
cggggcagat ccaaccttac ctggggcagc aaactttgtc ctggtacctg ggcccagctc   1885
gccagggatg tgggcagagc accagcctgg gcatcaggaa gccaagtttc aaggactgtc   1945
tttgagtctg tctgtatgac cttgggcctg ccacttctca cagaccctag gtatccacag   2005
ctgtgacatg ggggcaagcg gctttgtttc agcctaaccc aggagcttag taaaaattgc   2065
ataagaccag ggggaagagt gtcagcgtgg ggtgggaatt cccgcggcct ccacctgctt   2125
gctaggggca ggatctcatt caggctgccc tggaagcacc tgcttggccc tgccaccttc   2185
ctccagggga gggccagatg gcatcctggc ttggggcggg tgggacctac ccaggctctg   2245
agactttact ggcctatgcc tgaggcctct tttcctttaa ctccctaaat tatgatgact   2305
ccaagtccaa gcccacccct cccaaagatt gggaggttcc gccgttccca gaggctcctc   2365
ctgcggtgct cccaagactt ccatagacca tctggaccag tagcccatcc cgcagttttc   2425
ttgggggcag aggaaaacgc ttctttctcc tccagctgaa tcagctggat cccagtgtcc   2485
tggctgtttg gtgattgggc aagattgaat ttgcccaggt aggcgtgaga gtgtgggttt   2545
taaattcgaa gctcaggcca tagtttcaga gaatcaccct taccccagac cttcatgaga   2605
cagtgctcat gaagccagtg cgtttcccag aacgaacact aggcggcacc gttggtccac   2665
actcagaggc ccttggcgcc aagactgcat ctagaatcgc tcaaacacct gtttgcagac   2725
cccatgcacc agctggaggg gccgtaactg caggactgcg cctactgagt gacccatttc   2785
ctccaggagg aaaggcaaga cacgcttaca cggccatttg tctcttttcc caatgcggcg   2845
gtgcactttc gctcttgggg gctgcacccc agacatagct ggcaccagag cagggtgctc   2905
aggtggtggg tgctcagggc cctgccccag gccactgggc cgttttgatg acctcgaagg   2965
tcacaggcag aaaataggag caggatttcc cctggggaaa agttctcctg ggacatcttc   3025
tgctcttctg tacatttcta gatgcaaata actccttcac caggcagtga gtggcgtagg   3085
ctctggagcc aggctgcctg ggctccaatg ccagctctgc cacttgctag ctgtgagact   3145
gtggacaaac cactcagcct ctgtgtgcct cagttttcct atttgtaaaa tagaggccat   3205
agtggtacct attttgaaga ctaagtaaaa gaattcaaat aaagagactt ggc           3258
```

<210> 29
<211> 528

<212> PRT
<213> Homo sapiens

<400> 29

```
Met Asp His Cys Gly Ala Leu Phe Leu Cys Leu Cys Leu Leu Thr Leu
 1               5                  10                  15
Gln Asn Ala Thr Thr Glu Thr Trp Glu Glu Leu Leu Ser Tyr Met Glu
            20                  25                  30
Asn Met Gln Val Ser Arg Gly Arg Ser Ser Val Phe Ser Ser Arg Gln
        35                  40                  45
Leu His Gln Leu Glu Gln Met Leu Leu Asn Thr Ser Phe Pro Gly Tyr
    50                  55                  60
Asn Leu Thr Leu Gln Thr Pro Thr Ile Gln Ser Leu Ala Phe Lys Leu
65                  70                  75                  80
Ser Cys Asp Phe Ser Gly Leu Ser Leu Thr Ser Ala Thr Leu Lys Arg
                85                  90                  95
Val Pro Gln Ala Gly Gly Gln His Ala Arg Gly Gln His Ala Met Gln
            100                 105                 110
Phe Pro Ala Glu Leu Thr Arg Asp Ala Cys Lys Thr Arg Pro Arg Glu
        115                 120                 125
Leu Arg Leu Ile Cys Ile Tyr Phe Ser Asn Thr His Phe Phe Lys Asp
    130                 135                 140
Glu Asn Asn Ser Ser Leu Leu Asn Asn Tyr Val Leu Gly Ala Gln Leu
145                 150                 155                 160
Ser His Gly His Val Asn Asn Leu Arg Asp Pro Val Asn Ile Ser Phe
                165                 170                 175
Trp His Asn Gln Ser Leu Glu Gly Tyr Thr Leu Thr Cys Val Phe Trp
            180                 185                 190
Lys Glu Gly Ala Arg Lys Gln Pro Trp Gly Gly Trp Ser Pro Glu Gly
        195                 200                 205
Cys Arg Thr Glu Gln Pro Ser His Ser Gln Val Leu Cys Arg Cys Asn
    210                 215                 220
His Leu Thr Tyr Phe Ala Val Leu Met Gln Leu Ser Pro Ala Leu Val
225                 230                 235                 240
Pro Ala Glu Leu Leu Ala Pro Leu Thr Tyr Ile Ser Leu Val Gly Cys
                245                 250                 255
Ser Ile Ser Ile Val Ala Ser Leu Ile Thr Val Leu Leu His Phe His
            260                 265                 270
Phe Arg Lys Gln Ser Asp Ser Leu Thr Arg Ile His Met Asn Leu His
        275                 280                 285
Ala Ser Val Leu Leu Leu Asn Ile Ala Phe Leu Leu Ser Pro Ala Phe
    290                 295                 300
Ala Met Ser Pro Val Pro Gly Ser Ala Cys Thr Ala Leu Ala Ala Ala
305                 310                 315                 320
Leu His Tyr Ala Leu Leu Ser Cys Leu Thr Trp Met Ala Ile Glu Gly
                325                 330                 335
Phe Asn Leu Tyr Leu Leu Leu Gly Arg Val Tyr Asn Ile Tyr Ile Arg
            340                 345                 350
Arg Tyr Val Phe Lys Leu Gly Val Leu Gly Trp Gly Ala Pro Ala Leu
        355                 360                 365
Leu Val Leu Leu Ser Leu Ser Val Lys Ser Ser Val Tyr Gly Pro Cys
    370                 375                 380
Thr Ile Pro Val Phe Asp Ser Trp Glu Asn Gly Thr Gly Phe Gln Asn
385                 390                 395                 400
Met Ser Ile Cys Trp Val Arg Ser Pro Val Val His Ser Val Leu Val
                405                 410                 415
Met Gly Tyr Gly Gly Leu Thr Ser Leu Phe Asn Leu Val Val Leu Ala
            420                 425                 430
```

```
Trp Ala Leu Trp Thr Leu Arg Arg Leu Arg Glu Arg Ala Asp Ala Pro
        435                 440                 445
Ser Val Arg Ala Cys His Asp Thr Val Thr Val Leu Gly Leu Thr Val
    450                 455                 460
Leu Leu Gly Thr Thr Trp Ala Leu Ala Phe Phe Ser Phe Gly Val Phe
465                 470                 475                 480
Leu Leu Pro Gln Leu Phe Leu Phe Thr Ile Leu Asn Ser Leu Tyr Gly
                485                 490                 495
Phe Phe Leu Phe Leu Trp Phe Cys Ser Gln Arg Cys Arg Ser Glu Ala
            500                 505                 510
Glu Ala Lys Ala Gln Ile Glu Ala Phe Ser Ser Ser Gln Thr Thr Gln
        515                 520                 525
```

<210> 30
<211> 1584
<212> DNA
<213> Homo sapiens

<400> 30

```
atggatcact gtggtgccct tttcctgtgc ctgtgccttc tgactttgca gaatgcaaca      60
acagagacat gggaagaact cctgagctac atggagaata tgcaggtgtc caggggccgg     120
agctcagttt tttcctctcg tcaactccac cagctggagc agatgctact gaacaccagc     180
ttcccaggct acaacctgac cttgcagaca cccaccatcc agtctctggc cttcaagctg     240
agctgtgact tctctggcct ctcgctgacc agtgccactc tgaagcgggt gccccaggca     300
ggaggtcagc atgcccgggg tcagcacgcc atgcagttcc ccgccgagct gacccgggac     360
gcctgcaaga cccgccccag ggagctgcgg ctcatctgta tctacttctc caacacccac     420
tttttcaagg atgaaaacaa ctcatctctg ctgaataact acgtcctggg ggcccagctg     480
agtcatgggc acgtgaacaa cctcagggat cctgtgaaca tcagcttctg gcacaaccaa     540
agcctggaag gctacaccct gacctgtgtc ttctggaagg agggagccag gaaacagccc     600
tgggggggct ggagccctga gggctgtcgt acagagcagc cctcccactc tcaggtgctc     660
tgccgctgca accacctcac ctactttgct gttctcatgc aactctcccc agccctggtc     720
cctgcagagt tgctggcacc tcttacgtac atctccctcg tgggctgcag catctccatc     780
gtggcctcgc tgatcacagt cctgctgcac ttccatttca ggaagcagag tgactcctta     840
acacgcatcc acatgaacct gcatgcctcc gtgctgctcc tgaacatcgc cttcctgctg     900
agccccgcat tcgcaatgtc tcctgtgccc gggtcagcat gcacggctct ggccgctgcc     960
ctgcactacg cgctgctcag ctgcctcacc tggatggcca tcgagggctt caacctctac    1020
ctcctcctcg ggcgtgtcta caacatctac atccgcagat atgtgttcaa gcttggtgtg    1080
ctaggctggg gggccccagc cctcctggtg ctgctttccc tctctgtcaa gagctcggta    1140
tacggaccct gcacaatccc cgtcttcgac agctgggaga atggcacagg cttccagaac    1200
atgtccatat gctgggtgcg gagccccgtg gtgcacagtg tcctggtcat gggctacggc    1260
ggcctcacgt ccctcttcaa cctggtggtg ctggcctggg cgctgtggac cctgcgcagg    1320
ctgcgggagc gggcggatgc accaagtgtc agggcctgcc atgacactgt cactgtgctg    1380
ggcctcaccg tgctgctggg aaccacctgg gccttggcct tcttttcttt tggcgtcttc    1440
ctgctgcccc agctgttcct cttcaccatc ttaaactcgc tctacggttt cttcctttc     1500
ctgtggttct gctcccagcg gtgccgctca gaagcagagg ccaaggcaca gatagaggcc    1560
ttcagctcct cccaaacaac acag                                           1584
```

<210> 31
<211> 63
<212> PRT
<213> Homo sapiens

<400> 31

```
Leu Lys Ser Pro Glu Gly Lys Ser Arg Lys Asn Pro Ala Arg Thr Cys
 1               5                  10                  15
Lys Asp Leu Phe Leu Cys His Pro Glu Phe Lys Ser Gly Glu Tyr Trp
            20                  25                  30
Ile Asp Pro Asn Gln Gly Cys Ile Lys Asp Ala Ile Lys Val Phe Cys
```

```
        35                  40                  45
Asn Lys Arg Phe Glu Thr Gly Val Gly Glu Thr Cys Ile Ser Pro
    50                  55                  60

      <210> 32
      <211> 25
      <212> PRT
      <213> Homo sapiens

      <400> 32
Ile Ser Asn Val Gln Thr Phe Leu Arg Leu Leu Ser Thr Glu Ala Ser
 1               5                  10                  15
Gln Asn Ile Thr Tyr His Cys Lys Asn
            20                  25

      <210> 33
      <211> 33
      <212> PRT
      <213> Homo sapiens

      <400> 33
Thr Val Leu Gly Glu Asp Gly Cys Ser Ser Arg Thr Gly Glu Trp Gly
 1               5                  10                  15
Lys Thr Val Ile Glu Tyr Glu Thr Lys Lys Thr Thr Arg Leu Pro Ile
            20                  25                  30
Val


      <210> 34
      <211> 65
      <212> PRT
      <213> Homo sapiens

      <400> 34
Ile Asn Thr Ile Lys Asn Pro Leu Gly Thr Arg Asp Asn Pro Ala Arg
 1               5                  10                  15
Ile Cys Lys Asp Leu Leu Asn Cys Glu Gln Lys Val Ser Asp Gly Lys
            20                  25                  30
Tyr Trp Ile Asp Pro Asn Leu Gly Cys Pro Ser Asp Ala Ile Glu Val
        35                  40                  45
Phe Ile Asn Thr Cys Asn Phe Ser Ala Gly Gly Gln Thr Cys Leu Pro
    50                  55                  60
Pro
65

      <210> 35
      <211> 26
      <212> PRT
      <213> Homo sapiens

      <400> 35
Val Gly Lys Val Gln Met Asn Phe Leu His Leu Leu Ser Ser Glu Ala
 1               5                  10                  15
Thr His Ile Ile Thr Ile His Cys Leu Asn
            20                  25

      <210> 36
      <211> 32
```

<212> PRT
<213> Homo sapiens

<400> 36
Lys Val Leu Ser Asp Asp Cys Lys Ile Gln Asp Gly Ser Trp His Lys
1 5 10 15
Ala Thr Phe Leu Phe His Thr Gln Glu Pro Asn Gln Leu Pro Val Ile
20 25 30

<210> 37
<211> 31
<212> PRT
<213> Homo sapiens

<400> 37
Gly Glu Ser Val Thr Leu Thr Cys Ser Val Ser Gly Phe Gly Pro Pro
1 5 10 15
Pro Val Thr Trp Leu Arg Asn Gly Lys Leu Ser Leu Thr Ile Ser
20 25 30

<210> 38
<211> 57
<212> PRT
<213> Homo sapiens

<400> 38
Gly Arg Thr Val Arg Leu Gln Cys Pro Val Glu Gly Asp Pro Pro Pro
1 5 10 15
Thr Met Trp Thr Lys Asp Gly Arg Thr Ile His Ser Gly Trp Ser Arg
20 25 30
Phe Arg Val Leu Pro Gln Gly Leu Lys Val Lys Gln Val Glu Arg Glu
35 40 45
Asp Ala Gly Val Tyr Val Cys Lys Ala
50 55

<210> 39
<211> 59
<212> PRT
<213> Homo sapiens

<400> 39
Gly Ser Ser Val Arg Leu Lys Cys Val Ala Ser Gly His Pro Arg Pro
1 5 10 15
Asp Ile Thr Trp Met Lys Asp Asp Gln Ala Leu Thr Arg Pro Glu Ala
20 25 30
Ala Glu Pro Arg Lys Lys Lys Trp Thr Leu Ser Leu Lys Asn Leu Arg
35 40 45
Pro Glu Asp Ser Gly Lys Tyr Thr Cys Arg Val
50 55

<210> 40
<211> 79
<212> PRT
<213> Homo sapiens

<400> 40
Gly Gly Thr Thr Ser Phe Gln Cys Lys Val Arg Ser Asp Val Lys Pro
1 5 10 15

```
Val Ile Gln Trp Leu Lys Arg Val Glu Tyr Gly Ala Glu Gly Arg His
            20                  25                  30
Asn Ser Thr Ile Asp Val Gly Gly Gln Lys Phe Val Val Leu Pro Thr
        35                  40                  45
Gly Asp Val Trp Ser Arg Pro Asp Gly Ser Tyr Asn Lys Leu Leu Ile
    50                  55                  60
Thr Arg Ala Arg Gln Asp Asp Ala Gly Met Tyr Ile Cys Leu Gly
65                  70                  75
```

<210> 41
<211> 78
<212> PRT
<213> Homo sapiens

<400> 41

```
Arg Gly Ser Leu Thr Val Gln Cys Val Tyr Arg Ser Gly Trp Glu Thr
 1               5                  10                  15
Tyr Leu Lys Trp Trp Cys Arg Gly Ala Ile Trp Arg Asp Cys Lys Ile
            20                  25                  30
Leu Val Lys Thr Ser Gly Ser Glu Gln Glu Val Lys Arg Asp Arg Val
        35                  40                  45
Ser Ile Lys Asp Asn Gln Lys Asn Arg Thr Phe Thr Val Thr Met Glu
    50                  55                  60
Asp Leu Met Lys Thr Asp Ala Asp Thr Tyr Trp Cys Gly Ile
65                  70                  75
```

<210> 42
<211> 10
<212> PRT
<213> Homo sapiens

<400> 42

```
Val Phe Val Leu Gly Thr Leu Gly Ile Phe
 1               5                  10
```

<210> 43
<211> 10
<212> PRT
<213> Homo sapiens

<400> 43

```
Val Phe Ile Leu Gly Thr Leu Leu Leu Trp
 1               5                  10
```

<210> 44
<211> 116
<212> PRT
<213> Homo sapiens

<400> 44

```
Cys Gly Gly Thr Leu Asp Leu Thr Glu Ser Ser Gly Ser Ile Ser Ser
 1               5                  10                  15
Pro Asn Tyr Pro Asn Arg Ser Asp Tyr Pro Pro Asn Lys Glu Cys Val
            20                  25                  30
Trp Arg Ile Arg Ala Pro Pro Gly Tyr Arg Val Val Glu Leu Thr Phe
        35                  40                  45
Gln Asp Phe Asp Leu Glu Asp His Asp Gly Ala Pro Cys Arg Tyr Asp
    50                  55                  60
```

```
Tyr Val Glu Ile Arg Asp Gly Asp Pro Ser Ser Pro Leu Leu Gly Arg
65                  70                  75                  80
Phe Cys Gly Ser Gly Lys Pro Glu Asp Ile Arg Ser Thr Ser Asn Arg
                85                  90                  95
Met Leu Ile Lys Phe Val Ser Asp Ala Ser Val Ser Lys Arg Gly Phe
            100                 105                 110
Lys Ala Thr Tyr
        115
```

<210> 45
<211> 97
<212> PRT
<213> Homo sapiens

<400> 45

```
Gly Ser Val Leu Leu Ala Gln Glu Leu Pro Gln Gln Leu Thr Ser Pro
 1               5                  10                  15
Gly Tyr Pro Glu Pro Tyr Gly Lys Gly Gln Glu Ser Ser Thr Asp Ile
            20                  25                  30
Lys Ala Pro Glu Gly Phe Ala Val Arg Leu Val Phe Gln Asp Phe Asp
        35                  40                  45
Leu Glu Pro Ser Gln Asp Cys Ala Gly Asp Ser Val Thr Val Ser Trp
    50                  55                  60
Gly Trp Gly Gly Ser Arg Gln Asp Cys Gly Gln Gly Asp Ser Arg Gly
65                  70                  75                  80
Cys Gly Lys Trp Arg Cys Pro Glu Ser Pro Ile Trp Arg Arg Asp Glu
                85                  90                  95
Phe
```

<210> 46
<211> 45
<212> PRT
<213> Homo sapiens

<400> 46

```
Cys Ala Pro Asn Asn Pro Cys Ser Asn Gly Gly Thr Cys Val Asn Thr
 1               5                  10                  15
Pro Gly Gly Ser Ser Asp Asn Phe Gly Gly Tyr Thr Cys Glu Cys Pro
            20                  25                  30
Pro Gly Asp Tyr Tyr Leu Ser Tyr Thr Gly Lys Arg Cys
        35                  40                  45
```

<210> 47
<211> 67
<212> PRT
<213> Homo sapiens

<400> 47

```
Trp Ser Thr Asp Lys His Ile Gly Gly Arg Thr Ser Leu Gly Phe Asn
 1               5                  10                  15
Leu Glu Tyr Arg Ile Arg Val Thr Cys Asp Glu Asn Tyr Tyr Gly Glu
            20                  25                  30
Gly Cys Asn Lys Phe Cys Arg Pro Arg Asp Asp Ala Phe Gly His Tyr
        35                  40                  45
Thr Cys Asp Glu Asn Gly Asn Lys Leu Cys Leu Glu Gly Trp Lys Gly
    50                  55                  60
Glu Tyr Cys
```

65

<210> 48
<211> 59
<212> PRT
<213> Homo sapiens

<400> 48
Cys Asp Cys Asn Pro His Gly Ser Leu Ser Asp Asp Thr Cys Asp Ser
1 5 10 15
Asp Asp Glu Leu Phe Gly Glu Glu Thr Gly Gln Cys Leu Lys Cys Lys
20 25 30
Pro Asn Val Thr Gly Arg Arg Cys Asp Arg Cys Lys Pro Gly Tyr Tyr
35 40 45
Gly Leu Pro Ser Gly Asp Pro Gln Gln Gly Cys
50 55

<210> 49
<211> 31
<212> PRT
<213> Homo sapiens

<400> 49
Cys Val Pro Leu Cys Ala Gln Glu Cys Val His Gly Arg Cys Val Ala
1 5 10 15
Pro Asn Gln Cys Gln Cys Val Pro Gly Trp Arg Gly Asp Asp Cys
20 25 30

<210> 50
<211> 30
<212> PRT
<213> Homo sapiens

<400> 50
Cys Gln Phe Arg Cys Gln Cys His Gly Ala Pro Cys Asp Pro Gln Thr
1 5 10 15
Gly Ala Cys Phe Cys Pro Ala Glu Arg Thr Gly Pro Ser Cys
20 25 30

<210> 51
<211> 31
<212> PRT
<213> Homo sapiens

<400> 51
Cys Pro Ser Thr His Pro Cys Gln Asn Gly Gly Val Phe Gln Thr Pro
1 5 10 15
Gln Gly Ser Cys Ser Cys Pro Pro Gly Trp Met Gly Thr Ile Cys
20 25 30

<210> 52
<211> 31
<212> PRT
<213> Homo sapiens

<400> 52
Cys Ser Gln Glu Cys Arg Cys His Asn Gly Gly Leu Cys Asp Arg Phe
1 5 10 15

```
Thr Gly Gln Cys Arg Cys Ala Pro Gly Tyr Thr Gly Asp Arg Cys
            20                  25                  30

      <210> 53
      <211> 31
      <212> PRT
      <213> Homo sapiens

      <400> 53
Cys Ala Glu Thr Cys Asp Cys Ala Pro Asp Ala Arg Cys Phe Pro Ala
 1               5                  10                  15
Asn Gly Ala Cys Leu Cys Glu His Gly Phe Thr Gly Asp Arg Cys
            20                  25                  30

      <210> 54
      <211> 27
      <212> PRT
      <213> Homo sapiens

      <400> 54
Cys Asp Arg Glu His Ser Leu Ser Cys His Pro Met Asn Gly Glu Cys
 1               5                  10                  15
Ser Cys Leu Pro Gly Trp Ala Gly Leu His Cys
            20                  25

      <210> 55
      <211> 31
      <212> PRT
      <213> Homo sapiens

      <400> 55
Cys Gln Glu His Cys Leu Cys Leu His Gly Gly Val Cys Gln Ala Thr
 1               5                  10                  15
Ser Gly Leu Cys Gln Cys Ala Pro Gly Tyr Thr Gly Pro His Cys
            20                  25                  30

      <210> 56
      <211> 31
      <212> PRT
      <213> Homo sapiens

      <400> 56
Cys Ser Ala Arg Cys Ser Cys Glu Asn Ala Ile Ala Cys Ser Pro Ile
 1               5                  10                  15
Asp Gly Glu Cys Val Cys Lys Glu Gly Trp Gln Arg Gly Asn Cys
            20                  25                  30

      <210> 57
      <211> 31
      <212> PRT
      <213> Homo sapiens

      <400> 57
Cys Asn Ala Ser Cys Gln Cys Ala His Glu Ala Val Cys Ser Pro Gln
 1               5                  10                  15
Thr Gly Ala Cys Thr Cys Thr Pro Gly Trp His Gly Ala His Cys
            20                  25                  30
```

<210> 58
<211> 31
<212> PRT
<213> Homo sapiens

<400> 58
Cys Ala Ser Arg Cys Asp Cys Asp His Ser Asp Gly Cys Asp Pro Val
1 5 10 15
His Gly Arg Cys Gln Cys Gln Ala Gly Trp Met Gly Ala Arg Cys
20 25 30

<210> 59
<211> 31
<212> PRT
<213> Homo sapiens

<400> 59
Cys Ser Asn Thr Cys Thr Cys Lys Asn Gly Gly Thr Cys Leu Pro Glu
1 5 10 15
Asn Gly Asn Cys Val Cys Ala Pro Gly Phe Arg Gly Pro Ser Cys
20 25 30

<210> 60
<211> 30
<212> PRT
<213> Homo sapiens

<400> 60
Cys Val Pro Cys Lys Cys Ala Asn His Ser Phe Cys His Pro Ser Asn
1 5 10 15
Gly Thr Cys Tyr Cys Leu Ala Gly Trp Thr Gly Pro Asp Cys
20 25 30

<210> 61
<211> 31
<212> PRT
<213> Homo sapiens

<400> 61
Cys Ala Gln Thr Cys Gln Cys His His Gly Gly Thr Cys His Pro Gln
1 5 10 15
Asp Gly Ser Cys Ile Cys Pro Leu Gly Trp Thr Gly His His Cys
20 25 30

<210> 62
<211> 31
<212> PRT
<213> Homo sapiens

<400> 62
Cys Ser Gln Pro Cys Gln Cys Gly Pro Gly Glu Lys Cys His Pro Glu
1 5 10 15
Thr Gly Ala Cys Val Cys Pro Pro Gly His Ser Gly Ala Pro Cys
20 25 30

<210> 63
<211> 37
<212> PRT

<213> Homo sapiens

<400> 63
Gln Thr Gly Ala Cys Thr Cys Thr Pro Gly Trp His Gly Ala His Cys
1 5 10 15
Gln Leu Pro Cys Pro Lys Gly Gln Phe Gly Glu Gly Cys Ala Ser Arg
20 25 30
Cys Asp Cys Asp His
35

<210> 64
<211> 31
<212> PRT
<213> Mus musculus

<400> 64
Cys Ser Asn Thr Cys Thr Cys Lys Asn Gly Gly Thr Cys Val Ser Glu
1 5 10 15
Asn Gly Asn Cys Val Cys Ala Pro Gly Phe Arg Gly Pro Ser Cys
20 25 30

<210> 65
<211> 31
<212> PRT
<213> Mus musculus

<400> 65
Cys Val Gln Cys Lys Cys Asn Asn Asn His Ser Ser Cys His Pro Ser
1 5 10 15
Asp Gly Thr Cys Ser Cys Leu Ala Gly Trp Thr Gly Pro Asp Cys
20 25 30

<210> 66
<211> 31
<212> PRT
<213> Mus musculus

<400> 66
Cys Ser Gln Leu Cys Gln Cys His His Gly Gly Thr Cys His Pro Gln
1 5 10 15
Asp Gly Ser Cys Ile Cys Thr Pro Gly Trp Thr Gly Pro Asn Cys
20 25 30

<210> 67
<211> 31
<212> PRT
<213> Mus musculus

<400> 67
Cys Ser Gln Leu Cys Gln Cys Asp Leu Gly Glu Met Cys His Pro Glu
1 5 10 15
Thr Gly Ala Cys Val Cys Pro Pro Gly His Ser Gly Ala Asp Cys
20 25 30

<210> 68
<211> 35
<212> PRT
<213> Mus musculus

<400> 68
His Ala Ser Gly Asp Pro Val His Gly Gln Cys Arg Cys Gln Ala Gly
1 5 10 15
Trp Met Gly Thr Arg Cys His Leu Pro Cys Pro Glu Gly Phe Trp Gly
20 25 30
Ala Asn Cys
35

<210> 69
<211> 40
<212> PRT
<213> Mus musculus

<400> 69
Cys Thr Cys Lys Asn Gly Gly Thr Cys Val Ser Glu Asn Gly Asn Cys
1 5 10 15
Val Cys Ala Pro Gly Phe Arg Gly Pro Ser Cys Gln Arg Pro Cys Pro
20 25 30
Pro Gly Arg Tyr Gly Lys Arg Cys
35 40

<210> 70
<211> 35
<212> PRT
<213> Mus musculus

<400> 70
Cys Lys Cys Asn Asn Asn His Ser Ser Cys His Pro Ser Asp Gly Thr
1 5 10 15
Cys Ser Cys Leu Ala Gly Trp Thr Gly Pro Asp Cys Ser Glu Ala Cys
20 25 30
Pro Pro Gly
35

<210> 71
<211> 34
<212> PRT
<213> Mus musculus

<400> 71
Cys Gln Cys His His Gly Gly Thr Cys His Pro Gln Asp Gly Ser Cys
1 5 10 15
Ile Cys Thr Pro Gly Trp Thr Gly Pro Asn Cys Leu Glu Gly Cys Pro
20 25 30
Pro Arg

<210> 72
<211> 58
<212> PRT
<213> Mus musculus

<400> 72
His Gly Gln Cys Arg Cys Gln Ala Gly Trp Met Gly Thr Arg Cys His
1 5 10 15
Leu Pro Cys Pro Glu Gly Phe Trp Gly Ala Asn Cys Ser Asn Thr Cys
20 25 30
Thr Cys Lys Asn Gly Gly Thr Cys Val Ser Glu Asn Gly Asn Cys Val

```
        35                  40                  45
Cys Ala Pro Gly Phe Arg Gly Pro Ser Cys
    50                  55
```

<210> 73
<211> 28
<212> PRT
<213> Rattus sp.

<400> 73

```
Glu Cys Arg Cys His Asn Gly Gly Leu Cys Asp Arg Phe Thr Gly Gln
 1               5                  10                  15
Cys His Cys Ala Pro Gly Tyr Ile Gly Asp Arg Cys
            20                  25
```

<210> 74
<211> 31
<212> PRT
<213> Rattus sp.

<400> 74

```
Cys Ala Glu Thr Cys Asp Cys Ala Pro Gly Ala Arg Cys Phe Pro Ala
 1               5                  10                  15
Asn Gly Ala Cys Leu Cys Glu His Gly Phe Thr Gly Asp Arg Cys
            20                  25                  30
```

<210> 75
<211> 33
<212> PRT
<213> Rattus sp.

<400> 75

```
Cys Gln Asp Pro Cys Thr Cys Asp Pro Glu His Ser Leu Ser Cys His
 1               5                  10                  15
Pro Met His Gly Glu Cys Ser Cys Gln Pro Gly Trp Ala Gly Leu His
            20                  25                  30
Cys
```

<210> 76
<211> 31
<212> PRT
<213> Rattus sp.

<400> 76

```
Cys Gln Glu His Cys Leu Cys Leu His Gly Gly Val Cys Leu Ala Asp
 1               5                  10                  15
Ser Gly Leu Cys Arg Cys Ala Pro Gly Tyr Thr Gly Pro His Cys
            20                  25                  30
```

<210> 77
<211> 31
<212> PRT
<213> Rattus sp.

<400> 77

```
Cys Ser Ser His Cys Ser Cys Glu Asn Ala Ile Ala Cys Ser Pro Val
 1               5                  10                  15
```

```
Asp Gly Thr Cys Ile Cys Lys Glu Gly Trp Gln Arg Gly Asn Cys
            20                  25                  30

      <210> 78
      <211> 31
      <212> PRT
      <213> Rattus sp.

      <400> 78
Cys Asn Ala Ser Cys Gln Cys Ala His Glu Gly Val Cys Ser Pro Gln
 1               5                  10                  15
Thr Gly Ala Cys Thr Cys Thr Pro Gly Trp Arg Gly Val His Cys
            20                  25                  30

      <210> 79
      <211> 31
      <212> PRT
      <213> Rattus sp.

      <400> 79
Cys Ala Ser Val Cys Asp Cys Asp His Ser Asp Gly Cys Asp Pro Val
 1               5                  10                  15
His Gly His Cys Arg Cys Gln Ala Gly Trp Met Gly Thr Arg Cys
            20                  25                  30

      <210> 80
      <211> 31
      <212> PRT
      <213> Rattus sp.

      <400> 80
Cys Ser Asn Ala Cys Thr Cys Lys Asn Gly Gly Thr Cys Val Pro Glu
 1               5                  10                  15
Asn Gly Asn Cys Val Cys Ala Pro Gly Phe Arg Gly Pro Ser Cys
            20                  25                  30

      <210> 81
      <211> 30
      <212> PRT
      <213> Rattus sp.

      <400> 81
Cys Val Pro Cys Lys Cys Asn Asn His Ser Ser Cys His Pro Ser Asp
 1               5                  10                  15
Gly Thr Cys Ser Cys Leu Ala Gly Trp Thr Gly Pro Asp Cys
            20                  25                  30

      <210> 82
      <211> 31
      <212> PRT
      <213> Rattus sp.

      <400> 82
Cys Ser Gln Pro Cys Gln Cys His His Gly Ala Thr Cys His Pro Gln
 1               5                  10                  15
Asp Gly Ser Cys Val Cys Ile Pro Gly Trp Thr Gly Pro Asn Cys
            20                  25                  30
```

<210> 83
<211> 31
<212> PRT
<213> Rattus sp.

<400> 83
Cys Ser Gln Leu Cys Gln Cys Asp Pro Gly Glu Met Cys His Pro Glu
1 5 10 15
Thr Gly Ala Cys Val Cys Pro Pro Gly His Ser Gly Ala His Cys
20 25 30

<210> 84
<211> 40
<212> PRT
<213> Rattus sp.

<400> 84
Cys Arg Cys His Asn Gly Gly Leu Cys Asp Arg Phe Thr Gly Gln Cys
1 5 10 15
His Cys Ala Pro Gly Tyr Ile Gly Asp Arg Cys Arg Glu Glu Cys Pro
20 25 30
Val Gly Arg Phe Gly Gln Asp Cys
35 40

<210> 85
<211> 39
<212> PRT
<213> Rattus sp.

<400> 85
Cys Asp Cys Ala Pro Gly Ala Arg Cys Phe Pro Ala Asn Gly Ala Cys
1 5 10 15
Leu Cys Glu His Gly Phe Thr Gly Asp Arg Cys Thr Glu Arg Leu Cys
20 25 30
Pro Asp Gly Tyr Gly Leu Cys
35

<210> 86
<211> 42
<212> PRT
<213> Rattus sp.

<400> 86
Cys Thr Cys Asp Pro Glu His Ser Leu Ser Cys His Pro Met His Gly
1 5 10 15
Glu Cys Ser Cys Gln Pro Gly Trp Ala Gly Leu His Cys Asn Glu Ser
20 25 30
Cys Pro Gln Asp Thr His Gly Ala Gly Cys
35 40

<210> 87
<211> 40
<212> PRT
<213> Rattus sp.

<400> 87
Cys Leu Cys Leu His Gly Gly Val Cys Leu Ala Asp Ser Gly Leu Cys
1 5 10 15

Arg Cys Ala Pro Gly Tyr Thr Gly Pro His Cys Ala Asn Leu Cys Pro
20 25 30
Pro Asn Thr Tyr Gly Ile Asn Cys
35 40

<210> 88
<211> 40
<212> PRT
<213> Rattus sp.

<400> 88
Cys Ser Cys Glu Asn Ala Ile Ala Cys Ser Pro Val Asp Gly Thr Cys
1 5 10 15
Ile Cys Lys Glu Gly Trp Gln Arg Gly Asn Cys Ser Val Pro Cys Pro
20 25 30
Pro Gly Thr Trp Gly Phe Ser Cys
35 40

<210> 89
<211> 40
<212> PRT
<213> Rattus sp.

<400> 89
Cys Gln Cys Ala His Glu Gly Val Cys Ser Pro Gln Thr Gly Ala Cys
1 5 10 15
Thr Cys Thr Pro Gly Trp Arg Gly Val His Cys Gln Leu Pro Cys Pro
20 25 30
Lys Gly Gln Phe Gly Glu Gly Cys
35 40

<210> 90
<211> 40
<212> PRT
<213> Rattus sp.

<400> 90
Cys Asp Cys Asp His Ser Asp Gly Cys Asp Pro Val His Gly His Cys
1 5 10 15
Arg Cys Gln Ala Gly Trp Met Gly Thr Arg Cys His Leu Pro Cys Pro
20 25 30
Glu Gly Phe Trp Gly Ala Asn Cys
35 40

<210> 91
<211> 40
<212> PRT
<213> Rattus sp.

<400> 91
Cys Thr Cys Lys Asn Gly Gly Thr Cys Val Pro Glu Asn Gly Asn Cys
1 5 10 15
Val Cys Ala Pro Gly Phe Arg Gly Pro Ser Cys Gln Arg Pro Cys Pro
20 25 30
Pro Gly Arg Tyr Gly Lys Arg Cys
35 40

<210> 92

<211> 40
<212> PRT
<213> Rattus sp.

<400> 92
Cys Lys Cys Asn Asn His Ser Ser Cys His Pro Ser Asp Gly Thr Cys
1 5 10 15
Ser Cys Leu Ala Gly Trp Thr Gly Pro Asp Cys Ser Glu Ser Cys Pro
20 25 30
Pro Gly His Trp Gly Leu Lys Cys
35 40

<210> 93
<211> 40
<212> PRT
<213> Rattus sp.

<400> 93
Cys Gln Cys His His Gly Ala Thr Cys His Pro Gln Asp Gly Ser Cys
1 5 10 15
Val Cys Ile Pro Gly Trp Thr Gly Pro Asn Cys Ser Glu Gly Cys Pro
20 25 30
Ser Arg Met Phe Gly Val Asn Cys
35 40

<210> 94
<211> 36
<212> PRT
<213> Rattus sp.

<400> 94
Cys Gln Cys Asp Pro Gly Glu Met Cys His Pro Glu Thr Gly Ala Cys
1 5 10 15
Val Cys Pro Pro Gly His Ser Gly Ala His Cys Lys Val Gly Ser Gln
20 25 30
Glu Ser Phe Thr
35

<210> 95
<211> 64
<212> PRT
<213> Rattus sp.

<400> 95
Gly Val Cys Ser Pro Gln Thr Gly Ala Cys Thr Cys Thr Pro Gly Trp
1 5 10 15
Arg Gly Val His Cys Gln Leu Pro Cys Pro Lys Gly Gln Phe Gly Glu
20 25 30
Gly Cys Ala Ser Val Cys Asp Cys Asp His Ser Asp Gly Cys Asp Pro
35 40 45
Val His Gly His Cys Arg Cys Gln Ala Gly Trp Met Gly Thr Arg Cys
50 55 60

<210> 96
<211> 129
<212> PRT
<213> Homo sapiens

<400> 96
Gln Glu Ser Arg Ala Gln Lys Phe Leu Arg Gln His Ile Asp Ser Pro
1 5 10 15
Lys Thr Ser Ser Ser Asn Pro Asn Tyr Cys Asn Gln Met Met Asp Lys
20 25 30
Arg Arg Asn Met Thr Gln Gln Arg Cys Lys Pro Val Asn Thr Phe Val
35 40 45
His Glu Ser Leu Ala Asp Val Lys Ala Val Cys Ser Gln Lys Asn Val
50 55 60
Thr Cys Lys Asn Gly Gln Ser Lys Ser Ser Phe Gln Ile Thr Asp Cys
65 70 75 80
Arg Leu Thr Gly Gly Ser Gln Lys Tyr Pro Asn Cys Arg Tyr Arg Thr
85 90 95
Ser Ala Ser Thr Lys His Ile Ile Val Ala Cys Glu Gly Arg Asp Arg
100 105 110
Asp Asp Pro Tyr Tyr Asn Pro Tyr Val Pro Val His Phe Asp Ala Ser
115 120 125
Val

<210> 97
<211> 125
<212> PRT
<213> Homo sapiens

<400> 97
Gly Met Thr Ser Ser Gln Trp Phe Lys Ile Gln His Met Gln Pro Ser
1 5 10 15
Pro Gln Ala Cys Asn Ser Ala Met Lys Asn Ile Asn Lys His Thr Lys
20 25 30
Arg Cys Lys Asp Leu Asn Thr Phe Leu His Glu Pro Phe Ser Ser Val
35 40 45
Ala Ala Thr Cys Gln Thr Pro Lys Ile Ala Cys Lys Asn Gly Asp Lys
50 55 60
Asn Cys His Gln Ser His Gly Pro Val Ser Leu Thr Met Cys Lys Leu
65 70 75 80
Thr Ser Gly Lys Tyr Pro Asn Cys Arg Tyr Lys Glu Lys Arg Gln Asn
85 90 95
Lys Ser Tyr Val Val Ala Cys Lys Pro Pro Gln Lys Lys Asp Ser Gln
100 105 110
Gln Phe His Leu Val Pro Val His Leu Asp Arg Val Leu
115 120 125

<210> 98
<211> 411
<212> PRT
<213> Homo sapiens

<400> 98
Cys Asn Arg Thr Trp Asp Gly Ile Thr Cys Trp Pro Asp Thr Pro Pro
1 5 10 15
Gly Glu Leu Val Val Val Pro Cys Pro Lys Tyr Phe Tyr Gly Phe Ser
20 25 30
Ser Asp Gln Thr Asp Thr Thr Gly Asn Val Ser Arg Asn Cys Thr Glu
35 40 45
Asp Gly Ser Trp Ser Glu Pro Pro Pro Ser Asn Arg Thr Trp Arg Asn
50 55 60
Tyr Ser Ala Cys Gly Glu Asp Asp Pro Glu Glu Glu Ser Glu Lys Lys

```
65                  70                  75                  80
Lys Lys Tyr Tyr Leu Val Leu Lys Ile Ile Tyr Thr Val Gly Tyr Ser
                85                  90                  95
Leu Ser Leu Ala Ala Leu Leu Val Ala Val Val Ile Leu Leu Leu Phe
            100                 105                 110
Arg Lys Leu His Thr Leu Trp Pro Asp Asn Ala Asp Gly Ala Leu Glu
        115                 120                 125
Val Gly Ala Pro Trp Gly Ala Pro Phe Gln Val Arg Arg Ser Ile Arg
    130                 135                 140
Cys Thr Arg Asn Tyr Ile His Met Asn Leu Phe Leu Ser Phe Ile Leu
145                 150                 155                 160
Arg Ala Ala Ser Val Phe Ile Lys Asp Ala Val Leu Lys Ser Glu Val
                165                 170                 175
Ser Ser Asp Glu Pro Glu Arg Leu Ser Ser Arg Cys Ser Leu Ser Thr
            180                 185                 190
Gly Gln Val Val Val Gly Cys Lys Leu Leu Val Val Phe Gln Phe Gln
        195                 200                 205
Tyr Cys Val Met Thr Asn Phe Phe Trp Leu Leu Val Glu Gly Leu Tyr
    210                 215                 220
Leu His Thr Leu Leu Val Val Thr Phe Phe Ser Glu Arg Lys Tyr Leu
225                 230                 235                 240
Trp Trp Tyr Leu Leu Ile Gly Trp Gly Val Pro Leu Val Phe Val Thr
                245                 250                 255
Val Trp Ala Ile Val Arg Leu Leu Phe Glu Asp Thr Gly Cys Trp Asp
            260                 265                 270
Ser Asn Gly Leu Ala Met Phe Pro Glu Ala Lys Met Cys Ile Trp Met
        275                 280                 285
Ser Asp Asn Ser His Leu Trp Trp Ile Ile Lys Gly Pro Ile Leu Leu
    290                 295                 300
Ser Ile Leu Val Asn Phe Phe Leu Phe Ile Asn Ile Ile Arg Ile Leu
305                 310                 315                 320
Val Thr Lys Leu Arg Ala Ala Gln Thr Gly Glu Thr Asp Gln Arg Gln
                325                 330                 335
Tyr Ser Gln Tyr Arg Lys Leu Ala Lys Ser Thr Leu Leu Leu Ile Pro
            340                 345                 350
Leu Phe Gly Ile His Tyr Val Val Phe Ala Phe Arg Pro Ser Asn Asp
        355                 360                 365
Ala Arg Gly Val Leu Arg Lys Ile Lys Leu Tyr Phe Glu Leu Ser Leu
    370                 375                 380
Gly Ser Phe Gln Gly Phe Phe Val Ala Val Leu Tyr Cys Phe Leu Asn
385                 390                 395                 400
Gly Glu Val Gln Ala Glu Ile Arg Arg Arg Trp
                405                 410
```

<210> 99
<211> 328
<212> PRT
<213> Homo sapiens

<400> 99

```
Leu Thr Cys Val Phe Trp Lys Glu Gly Ala Arg Lys Gln Pro Trp Gly
 1               5                  10                  15
Gly Trp Ser Pro Glu Gly Cys Arg Thr Glu Gln Pro Ser His Ser Gln
            20                  25                  30
Val Leu Cys Arg Cys Asn His Leu Thr Tyr Phe Ala Val Leu Met Gln
        35                  40                  45
Leu Ser Pro Ala Leu Val Pro Ala Glu Leu Leu Ala Pro Leu Thr Tyr
    50                  55                  60
```

```
Ile Ser Leu Val Gly Cys Ser Ile Ser Ile Val Ala Ser Leu Ile Thr
65                  70                  75                  80
Val Leu Leu His Phe Arg Lys Gln Ser Asp Ser Leu Thr Arg Ile His
                85                  90                  95
Met Asn Leu His Ala Ser Val Leu Leu Leu Asn Ile Ala Phe Leu Leu
            100                 105                 110
Ser Pro Ala Phe Ala Met Ser Pro Val Pro Gly Ser Ala Cys Thr Ala
        115                 120                 125
Leu Ala Ala Ala Leu His Tyr Ala Leu Leu Ser Cys Leu Thr Trp Met
    130                 135                 140
Ala Ile Glu Gly Phe Asn Leu Tyr Leu Leu Leu Gly Arg Val Tyr Asn
145                 150                 155                 160
Ile Tyr Ile Arg Arg Tyr Val Phe Lys Leu Gly Val Leu Gly Trp Gly
                165                 170                 175
Ala Pro Ala Leu Leu Val Leu Leu Ser Leu Ser Val Lys Ser Ser Val
            180                 185                 190
Tyr Gly Pro Cys Thr Ile Pro Val Phe Asp Ser Trp Glu Asn Gly Thr
        195                 200                 205
Gly Phe Gln Asn Met Ser Ile Cys Trp Val Arg Ser Pro Val Val His
    210                 215                 220
Ser Val Leu Val Met Gly Tyr Gly Gly Leu Thr Ser Leu Phe Asn Leu
225                 230                 235                 240
Val Val Leu Ala Trp Ala Leu Trp Thr Leu Arg Arg Leu Arg Glu Arg
                245                 250                 255
Ala Asp Ala Pro Ser Val Arg Ala Cys His Asp Thr Val Thr Val Leu
            260                 265                 270
Gly Leu Thr Val Leu Leu Gly Thr Thr Trp Ala Leu Ala Phe Phe Ser
        275                 280                 285
Phe Gly Val Phe Leu Leu Pro Gln Leu Phe Leu Phe Thr Ile Leu Asn
    290                 295                 300
Ser Leu Tyr Gly Phe Phe Leu Phe Leu Trp Phe Cys Ser Gln Arg Cys
305                 310                 315                 320
Arg Ser Glu Ala Glu Ala Lys Ala
                325
```

<210> 100
<211> 150
<212> PRT
<213> Pan troglodytes

<400> 100

```
Met Val Leu Cys Phe Pro Leu Leu Leu Leu Leu Leu Val Leu Trp Gly
 1                  5                  10                  15
Pro Val Cys Pro Leu His Ala Trp Pro Lys Arg Leu Thr Lys Ala His
            20                  25                  30
Trp Phe Glu Ile Gln His Ile Gln Pro Ser Pro Leu Gln Cys Asn Arg
        35                  40                  45
Ala Met Ser Gly Ile Asn Asn Tyr Ala Gln His Cys Lys His Gln Asn
    50                  55                  60
Thr Phe Leu His Asp Ser Phe Gln Asn Val Ala Ala Val Cys Asp Leu
65                  70                  75                  80
Leu Ser Ile Val Cys Lys Asn Arg Arg His Asn Cys His Gln Ser Ser
                85                  90                  95
Lys Pro Val Asn Met Thr Asp Cys Arg Leu Thr Ser Gly Lys Tyr Pro
            100                 105                 110
Gln Cys Arg Tyr Ser Ala Ala Ala Gln Tyr Lys Phe Phe Ile Val Ala
        115                 120                 125
Cys Asp Pro Pro Gln Lys Ser Asp Pro Pro Tyr Lys Leu Val Pro Val
```

```
    130                 135                 140
His Leu Asp Ser Ile Leu
145                 150

      <210> 101
      <211> 24
      <212> PRT
      <213> Homo sapiens

      <400> 101
Met Thr Pro Ser Pro Leu Leu Leu Leu Leu Leu Pro Pro Leu Leu Leu
 1               5                  10                  15
Gly Ala Phe Pro Pro Ala Ala Ala
            20

      <210> 102
      <211> 480
      <212> PRT
      <213> Homo sapiens

      <400> 102
Ala Arg Gly Pro Pro Lys Met Ala Asp Lys Val Val Pro Arg Gln Val
 1               5                  10                  15
Ala Arg Leu Gly Arg Thr Val Arg Leu Gln Cys Pro Val Glu Gly Asp
            20                  25                  30
Pro Pro Pro Leu Thr Met Trp Thr Lys Asp Gly Arg Thr Ile His Ser
        35                  40                  45
Gly Trp Ser Arg Phe Arg Val Leu Pro Gln Gly Leu Lys Val Lys Gln
    50                  55                  60
Val Glu Arg Glu Asp Ala Gly Val Tyr Val Cys Lys Ala Thr Asn Gly
65                  70                  75                  80
Phe Gly Ser Leu Ser Val Asn Tyr Thr Leu Val Val Leu Asp Asp Ile
                85                  90                  95
Ser Pro Gly Lys Glu Ser Leu Gly Pro Asp Ser Ser Ser Gly Gly Gln
            100                 105                 110
Glu Asp Pro Ala Ser Gln Gln Trp Ala Arg Pro Arg Phe Thr Gln Pro
        115                 120                 125
Ser Lys Met Arg Arg Arg Val Ile Ala Arg Pro Val Gly Ser Ser Val
    130                 135                 140
Arg Leu Lys Cys Val Ala Ser Gly His Pro Arg Pro Asp Ile Thr Trp
145                 150                 155                 160
Met Lys Asp Asp Gln Ala Leu Thr Arg Pro Glu Ala Ala Glu Pro Arg
                165                 170                 175
Lys Lys Lys Trp Thr Leu Ser Leu Lys Asn Leu Arg Pro Glu Asp Ser
            180                 185                 190
Gly Lys Tyr Thr Cys Arg Val Ser Asn Arg Ala Gly Ala Ile Asn Ala
        195                 200                 205
Thr Tyr Lys Val Asp Val Ile Gln Arg Thr Arg Ser Lys Pro Val Leu
    210                 215                 220
Thr Gly Thr His Pro Val Asn Thr Thr Val Asp Phe Gly Gly Thr Thr
225                 230                 235                 240
Ser Phe Gln Cys Lys Val Arg Ser Asp Val Lys Pro Val Ile Gln Trp
                245                 250                 255
Leu Lys Arg Val Glu Tyr Gly Ala Glu Gly Arg His Asn Ser Thr Ile
            260                 265                 270
Asp Val Gly Gly Gln Lys Phe Val Val Leu Pro Thr Gly Asp Val Trp
        275                 280                 285
Ser Arg Pro Asp Gly Ser Tyr Leu Asn Lys Leu Leu Ile Thr Arg Ala
```

```
    290                 295                 300
Arg Gln Asp Asp Ala Gly Met Tyr Ile Cys Leu Gly Ala Asn Thr Met
305                 310                 315                 320
Gly Tyr Ser Phe Arg Ser Ala Phe Leu Thr Val Leu Pro Asp Pro Lys
                325                 330                 335
Pro Pro Gly Pro Pro Val Ala Ser Ser Ser Ser Ala Thr Ser Leu Pro
            340                 345                 350
Trp Pro Val Val Ile Gly Ile Pro Ala Gly Ala Val Phe Ile Leu Gly
        355                 360                 365
Thr Leu Leu Leu Trp Leu Cys Gln Ala Gln Lys Lys Pro Cys Thr Pro
    370                 375                 380
Ala Pro Ala Pro Pro Leu Pro Gly His Arg Pro Pro Gly Thr Ala Arg
385                 390                 395                 400
Asp Arg Ser Gly Asp Lys Asp Leu Pro Ser Leu Ala Ala Leu Ser Ala
                405                 410                 415
Gly Pro Gly Val Gly Leu Cys Glu Glu His Gly Ser Pro Ala Ala Pro
            420                 425                 430
Gln His Leu Leu Gly Pro Gly Pro Val Ala Gly Pro Lys Leu Tyr Pro
        435                 440                 445
Lys Leu Tyr Thr Asp Ile His Thr His Thr His Thr His Ser His Thr
    450                 455                 460
His Ser His Val Glu Gly Lys Val His Gln His Ile His Tyr Gln Cys
465                 470                 475                 480
```

<210> 103
<211> 350
<212> PRT
<213> Homo sapiens

<400> 103

```
Ala Arg Gly Pro Pro Lys Met Ala Asp Lys Val Val Pro Arg Gln Val
 1               5                  10                  15
Ala Arg Leu Gly Arg Thr Val Arg Leu Gln Cys Pro Val Glu Gly Asp
            20                  25                  30
Pro Pro Pro Leu Thr Met Trp Thr Lys Asp Gly Arg Thr Ile His Ser
        35                  40                  45
Gly Trp Ser Arg Phe Arg Val Leu Pro Gln Gly Leu Lys Val Lys Gln
    50                  55                  60
Val Glu Arg Glu Asp Ala Gly Val Tyr Val Cys Lys Ala Thr Asn Gly
65                  70                  75                  80
Phe Gly Ser Leu Ser Val Asn Tyr Thr Leu Val Val Leu Asp Asp Ile
                85                  90                  95
Ser Pro Gly Lys Glu Ser Leu Gly Pro Asp Ser Ser Ser Gly Gly Gln
            100                 105                 110
Glu Asp Pro Ala Ser Gln Gln Trp Ala Arg Pro Arg Phe Thr Gln Pro
        115                 120                 125
Ser Lys Met Arg Arg Arg Val Ile Ala Arg Pro Val Gly Ser Ser Val
    130                 135                 140
Arg Leu Lys Cys Val Ala Ser Gly His Pro Arg Pro Asp Ile Thr Trp
145                 150                 155                 160
Met Lys Asp Asp Gln Ala Leu Thr Arg Pro Glu Ala Ala Glu Pro Arg
                165                 170                 175
Lys Lys Lys Trp Thr Leu Ser Leu Lys Asn Leu Arg Pro Glu Asp Ser
            180                 185                 190
Gly Lys Tyr Thr Cys Arg Val Ser Asn Arg Ala Gly Ala Ile Asn Ala
        195                 200                 205
Thr Tyr Lys Val Asp Val Ile Gln Arg Thr Arg Ser Lys Pro Val Leu
    210                 215                 220
```

```
Thr Gly Thr His Pro Val Asn Thr Thr Val Asp Phe Gly Gly Thr Thr
225                 230                 235                 240
Ser Phe Gln Cys Lys Val Arg Ser Asp Val Lys Pro Val Ile Gln Trp
                245                 250                 255
Leu Lys Arg Val Glu Tyr Gly Ala Glu Gly Arg His Asn Ser Thr Ile
            260                 265                 270
Asp Val Gly Gly Gln Lys Phe Val Val Leu Pro Thr Gly Asp Val Trp
        275                 280                 285
Ser Arg Pro Asp Gly Ser Tyr Leu Asn Lys Leu Leu Ile Thr Arg Ala
    290                 295                 300
Arg Gln Asp Asp Ala Gly Met Tyr Ile Cys Leu Gly Ala Asn Thr Met
305                 310                 315                 320
Gly Tyr Ser Phe Arg Ser Ala Phe Leu Thr Val Leu Pro Asp Pro Lys
                325                 330                 335
Pro Pro Gly Pro Pro Val Ala Ser Ser Ser Ser Ala Thr Ser
            340                 345                 350
```

<210> 104
<211> 24
<212> PRT
<213> Homo sapiens

<400> 104

```
Leu Pro Trp Pro Val Val Ile Gly Ile Pro Ala Gly Ala Val Phe Ile
 1               5                  10                  15
Leu Gly Thr Leu Leu Leu Trp Leu
            20
```

<210> 105
<211> 106
<212> PRT
<213> Homo sapiens

<400> 105

```
Cys Gln Ala Gln Lys Lys Pro Cys Thr Pro Ala Pro Ala Pro Pro Leu
 1               5                  10                  15
Pro Gly His Arg Pro Pro Gly Thr Ala Arg Asp Arg Ser Gly Asp Lys
            20                  25                  30
Asp Leu Pro Ser Leu Ala Ala Leu Ser Ala Gly Pro Gly Val Gly Leu
        35                  40                  45
Cys Glu Glu His Gly Ser Pro Ala Ala Pro Gln His Leu Leu Gly Pro
    50                  55                  60
Gly Pro Val Ala Gly Pro Lys Leu Tyr Pro Lys Leu Tyr Thr Asp Ile
65                  70                  75                  80
His Thr His Thr His Thr His Ser His Thr His Ser His Val Glu Gly
                85                  90                  95
Lys Val His Gln His Ile His Tyr Gln Cys
            100                 105
```

<210> 106
<211> 208
<212> PRT
<213> Mus musculus

<400> 106

```
Arg Val Arg Pro Thr Gly Asp Val Trp Ser Arg Pro Asp Gly Ser Tyr
 1               5                  10                  15
Leu Asn Lys Leu Leu Ile Ser Arg Ala Arg Gln Asp Asp Ala Gly Met
```

```
                20                  25                  30
Tyr Ile Cys Leu Gly Ala Asn Thr Met Gly Tyr Ser Phe Arg Ser Ala
            35                  40                  45
Phe Leu Thr Val Leu Pro Asp Pro Lys Pro Pro Gly Pro Pro Met Ala
        50                  55                  60
Ser Ser Ser Ser Ser Thr Ser Leu Pro Trp Pro Val Val Ile Gly Ile
65                  70                  75                  80
Pro Ala Gly Ala Val Phe Ile Leu Gly Thr Val Leu Leu Trp Leu Cys
                85                  90                  95
Gln Thr Lys Lys Lys Pro Cys Ala Pro Ala Ser Thr Leu Pro Val Pro
            100                 105                 110
Gly His Arg Pro Pro Gly Thr Ser Arg Glu Arg Ser Gly Asp Lys Asp
        115                 120                 125
Leu Pro Ser Leu Ala Val Gly Ile Cys Glu Glu His Gly Ser Ala Met
    130                 135                 140
Ala Pro Gln His Ile Leu Ala Ser Gly Ser Thr Ala Gly Pro Lys Leu
145                 150                 155                 160
Tyr Pro Lys Leu Tyr Thr Asp Val His Thr His Thr His Thr His Thr
                165                 170                 175
Cys Thr His Thr Leu Ser Cys Trp Arg Ala Arg Phe Ile Asn Thr Ser
            180                 185                 190
Met Ser Thr Ile Ser Ala Lys Tyr Ser Glu Ser Pro Ser Thr Val Ser
        195                 200                 205

      <210> 107
      <211> 73
      <212> PRT
      <213> Mus musculus

      <400> 107
Arg Val Arg Pro Thr Gly Asp Val Trp Ser Arg Pro Asp Gly Ser Tyr
 1               5                  10                  15
Leu Asn Lys Leu Leu Ile Ser Arg Ala Arg Gln Asp Asp Ala Gly Met
            20                  25                  30
Tyr Ile Cys Leu Gly Ala Asn Thr Met Gly Tyr Ser Phe Arg Ser Ala
        35                  40                  45
Phe Leu Thr Val Leu Pro Asp Pro Lys Pro Pro Gly Pro Pro Met Ala
    50                  55                  60
Ser Ser Ser Ser Ser Thr Ser Leu Pro
65                  70

      <210> 108
      <211> 23
      <212> PRT
      <213> Mus musculus

      <400> 108
Trp Pro Val Val Ile Gly Ile Pro Ala Gly Ala Val Phe Ile Leu Gly
 1               5                  10                  15
Thr Val Leu Leu Trp Leu Cys
            20

      <210> 109
      <211> 112
      <212> PRT
      <213> Mus musculus

      <400> 109
```

```
Gln Thr Lys Lys Lys Pro Cys Ala Pro Ala Ser Thr Leu Pro Val Pro
 1               5                  10                  15
Gly His Arg Pro Pro Gly Thr Ser Arg Glu Arg Ser Gly Asp Lys Asp
            20                  25                  30
Leu Pro Ser Leu Ala Val Gly Ile Cys Glu Glu His Gly Ser Ala Met
        35                  40                  45
Ala Pro Gln His Ile Leu Ala Ser Gly Ser Thr Ala Gly Pro Lys Leu
    50                  55                  60
Tyr Pro Lys Leu Tyr Thr Asp Val His Thr His Thr His Thr His Thr
65                  70                  75                  80
Cys Thr His Thr Leu Ser Cys Trp Arg Ala Arg Phe Ile Asn Thr Ser
                85                  90                  95
Met Ser Thr Ile Ser Ala Lys Tyr Ser Glu Ser Pro Ser Thr Val Ser
            100                 105                 110
```

<210> 110
<211> 35
<212> PRT
<213> Homo sapiens

<400> 110

```
Met Pro Gly Pro Arg Val Trp Gly Lys Tyr Leu Trp Arg Ser Pro His
 1               5                  10                  15
Ser Lys Gly Cys Pro Gly Ala Met Trp Trp Leu Leu Leu Trp Gly Val
            20                  25                  30
Leu Gln Ala
        35
```

<210> 111
<211> 103
<212> PRT
<213> Homo sapiens

<400> 111

```
Cys Pro Thr Arg Gly Ser Val Leu Leu Ala Gln Glu Leu Pro Gln Gln
 1               5                  10                  15
Leu Thr Ser Pro Gly Tyr Pro Glu Pro Tyr Gly Lys Gly Gln Glu Ser
            20                  25                  30
Ser Thr Asp Ile Lys Ala Pro Glu Gly Phe Ala Val Arg Leu Val Phe
        35                  40                  45
Gln Asp Phe Asp Leu Glu Pro Ser Gln Asp Cys Ala Gly Asp Ser Val
    50                  55                  60
Thr Val Ser Trp Gly Trp Gly Gly Ser Arg Gln Asp Cys Gly Gln Gly
65                  70                  75                  80
Asp Ser Arg Gly Cys Gly Lys Trp Arg Cys Pro Glu Ser Pro Ile Trp
                85                  90                  95
Arg Arg Asp Glu Phe Ser Met
            100
```

<210> 112
<211> 20
<212> PRT
<213> Homo sapiens

<400> 112

```
Met Ser Pro Pro Leu Cys Pro Leu Leu Leu Leu Ala Val Gly Leu Arg
 1               5                  10                  15
Leu Ala Gly Thr
```

20

<210> 113
<211> 1030
<212> PRT
<213> Homo sapiens

<400> 113

```
Leu Asn Pro Ser Asp Pro Asn Thr Cys Ser Phe Trp Glu Ser Phe Thr
 1               5                  10                  15
Thr Thr Thr Lys Glu Ser His Ser Arg Pro Phe Ser Leu Leu Pro Ser
            20                  25                  30
Glu Pro Cys Glu Arg Pro Trp Glu Gly Pro His Thr Cys Pro Ser Pro
        35                  40                  45
Gln Thr Gln Arg Lys Leu Leu Ala Ser Arg Asp Ser Phe Cys Met Val
    50                  55                  60
Cys Val Gly Ala Gly Val Gln Trp Arg Asp Arg Ser Ala Leu Gln Pro
65                  70                  75                  80
Gln Thr Gly Asn Ala Leu Ser Met Arg Pro Gln Pro Arg Val Leu Ser
                85                  90                  95
Gly Ala Pro Ser Leu Ala Ser Pro Gly His Thr Val Val Val Lys Thr
            100                 105                 110
Asp His Arg Gln Arg Leu Gln Cys Cys His Gly Phe Tyr Glu Ser Arg
        115                 120                 125
Gly Phe Cys Val Pro Leu Cys Ala Gln Glu Cys Val His Gly Arg Cys
    130                 135                 140
Val Ala Pro Asn Gln Cys Gln Cys Val Pro Gly Trp Arg Gly Asp Asp
145                 150                 155                 160
Cys Ser Ser Ala Pro Asn Cys Leu Gln Pro Cys Thr Pro Gly Tyr Tyr
                165                 170                 175
Gly Pro Ala Cys Gln Phe Arg Cys Gln Cys His Gly Ala Pro Cys Asp
            180                 185                 190
Pro Gln Thr Gly Ala Cys Phe Cys Pro Ala Glu Arg Thr Gly Pro Ser
        195                 200                 205
Cys Asp Val Ser Cys Ser Gln Gly Thr Ser Gly Phe Phe Cys Pro Ser
    210                 215                 220
Thr His Pro Cys Gln Asn Gly Gly Val Phe Gln Thr Pro Gln Gly Ser
225                 230                 235                 240
Cys Ser Cys Pro Pro Gly Trp Met Gly Thr Ile Cys Ser Leu Pro Cys
                245                 250                 255
Pro Glu Gly Phe His Gly Pro Asn Cys Ser Gln Glu Cys Arg Cys His
            260                 265                 270
Asn Gly Gly Leu Cys Asp Arg Phe Thr Gly Gln Cys Arg Cys Ala Pro
        275                 280                 285
Gly Tyr Thr Gly Asp Arg Cys Arg Glu Glu Cys Pro Val Gly Arg Phe
    290                 295                 300
Gly Gln Asp Cys Ala Glu Thr Cys Asp Cys Ala Pro Asp Ala Arg Cys
305                 310                 315                 320
Phe Pro Ala Asn Gly Ala Cys Leu Cys Glu His Gly Phe Thr Gly Asp
                325                 330                 335
Arg Cys Thr Asp Arg Leu Cys Pro Asp Gly Phe Tyr Gly Leu Ser Cys
            340                 345                 350
Gln Ala Pro Cys Thr Cys Asp Arg Glu His Ser Leu Ser Cys His Pro
        355                 360                 365
Met Asn Gly Glu Cys Ser Cys Leu Pro Gly Trp Ala Gly Leu His Cys
    370                 375                 380
Asn Glu Ser Cys Pro Gln Asp Thr His Gly Pro Gly Cys Gln Glu His
385                 390                 395                 400
```

```
Cys Leu Cys Leu His Gly Gly Val Cys Gln Ala Thr Ser Gly Leu Cys
                405                 410                 415
Gln Cys Ala Pro Gly Tyr Thr Gly Pro His Cys Ala Ser Leu Cys Pro
            420                 425                 430
Pro Asp Thr Tyr Gly Val Asn Cys Ser Ala Arg Cys Ser Cys Glu Asn
        435                 440                 445
Ala Ile Ala Cys Ser Pro Ile Asp Gly Glu Cys Val Cys Lys Glu Gly
    450                 455                 460
Trp Gln Arg Gly Asn Cys Ser Val Pro Cys Pro Pro Gly Thr Trp Gly
465                 470                 475                 480
Phe Ser Cys Asn Ala Ser Cys Gln Cys Ala His Glu Ala Val Cys Ser
                485                 490                 495
Pro Gln Thr Gly Ala Cys Thr Cys Thr Pro Gly Trp His Gly Ala His
            500                 505                 510
Cys Gln Leu Pro Cys Pro Lys Gly Gln Phe Gly Glu Gly Cys Ala Ser
        515                 520                 525
Arg Cys Asp Cys Asp His Ser Asp Gly Cys Asp Pro Val His Gly Arg
    530                 535                 540
Cys Gln Cys Gln Ala Gly Trp Met Gly Ala Arg Cys His Leu Ser Cys
545                 550                 555                 560
Pro Glu Gly Leu Trp Gly Val Asn Cys Ser Asn Thr Cys Thr Cys Lys
                565                 570                 575
Asn Gly Gly Thr Cys Leu Pro Glu Asn Gly Asn Cys Val Cys Ala Pro
            580                 585                 590
Gly Phe Arg Gly Pro Ser Cys Gln Arg Ser Cys Gln Pro Gly Arg Tyr
        595                 600                 605
Gly Lys Arg Cys Val Pro Cys Lys Cys Ala Asn His Ser Phe Cys His
    610                 615                 620
Pro Ser Asn Gly Thr Cys Tyr Cys Leu Ala Gly Trp Thr Gly Pro Asp
625                 630                 635                 640
Cys Ser Gln Pro Cys Pro Pro Gly His Trp Gly Glu Asn Cys Ala Gln
                645                 650                 655
Thr Cys Gln Cys His His Gly Gly Thr Cys His Pro Gln Asp Gly Ser
            660                 665                 670
Cys Ile Cys Pro Leu Gly Trp Thr Gly His His Cys Leu Glu Gly Cys
        675                 680                 685
Pro Leu Gly Thr Phe Gly Ala Asn Cys Ser Gln Pro Cys Gln Cys Gly
    690                 695                 700
Pro Gly Glu Lys Cys His Pro Glu Thr Gly Ala Cys Val Cys Pro Pro
705                 710                 715                 720
Gly His Ser Gly Ala Pro Cys Arg Ile Gly Ile Gln Glu Pro Phe Thr
                725                 730                 735
Val Met Pro Thr Thr Pro Val Ala Tyr Asn Ser Leu Gly Ala Val Ile
            740                 745                 750
Gly Ile Ala Val Leu Gly Ser Leu Val Val Ala Leu Val Ala Leu Phe
        755                 760                 765
Ile Gly Tyr Arg His Trp Gln Lys Gly Lys Glu His His His Leu Ala
    770                 775                 780
Val Ala Tyr Ser Ser Gly Arg Leu Asp Gly Ser Glu Tyr Val Met Pro
785                 790                 795                 800
Asp Val Pro Pro Ser Tyr Ser His Tyr Tyr Ser Asn Pro Ser Tyr His
                805                 810                 815
Thr Leu Ser Gln Cys Ser Pro Asn Pro Pro Pro Pro Asn Lys Val Pro
            820                 825                 830
Gly Pro Leu Phe Ala Ser Leu Gln Asn Pro Glu Arg Pro Gly Gly Ala
        835                 840                 845
Gln Gly His Asp Asn His Thr Thr Leu Pro Ala Asp Trp Lys His Arg
    850                 855                 860
```

```
Arg Glu Pro Pro Pro Gly Pro Leu Asp Arg Gly Ser Ser Arg Leu Asp
865                 870                 875                 880
Arg Ser Tyr Ser Tyr Ser Tyr Ser Asn Gly Pro Gly Pro Phe Tyr Asp
            885                 890                 895
Lys Gly Leu Ile Ser Glu Glu Glu Leu Gly Ala Ser Val Ala Ser Leu
        900                 905                 910
Ser Ser Glu Asn Pro Tyr Ala Thr Ile Arg Asp Leu Pro Ser Leu Pro
    915                 920                 925
Gly Gly Pro Arg Glu Ser Ser Tyr Met Glu Met Lys Gly Pro Pro Ser
930                 935                 940
Gly Ser Ala Pro Arg Gln Pro Pro Gln Phe Trp Asp Ser Gln Arg Arg
945                 950                 955                 960
Arg Gln Pro Gln Pro Gln Arg Asp Ser Gly Thr Tyr Glu Gln Pro Ser
                965                 970                 975
Pro Leu Ile His Asp Arg Asp Ser Val Gly Ser Gln Pro Pro Leu Pro
            980                 985                 990
Pro Gly Leu Pro Pro Gly His Tyr Asp Ser Pro Lys Asn Ser His Ile
        995                 1000                1005
Pro Gly His Tyr Asp Leu Pro Pro Val Arg His Pro Pro Ser Pro Pro
    1010                1015                1020
Leu Arg Arg Gln Asp Arg
1025                1030
```

<210> 114
<211> 747
<212> PRT
<213> Homo sapiens

<400> 114

```
Leu Asn Pro Ser Asp Pro Asn Thr Cys Ser Phe Trp Glu Ser Phe Thr
 1               5                  10                  15
Thr Thr Thr Lys Glu Ser His Ser Arg Pro Phe Ser Leu Leu Pro Ser
            20                  25                  30
Glu Pro Cys Glu Arg Pro Trp Glu Gly Pro His Thr Cys Pro Ser Pro
        35                  40                  45
Gln Thr Gln Arg Lys Leu Leu Ala Ser Arg Asp Ser Phe Cys Met Val
    50                  55                  60
Cys Val Gly Ala Gly Val Gln Trp Arg Asp Arg Ser Ala Leu Gln Pro
65                  70                  75                  80
Gln Thr Gly Asn Ala Leu Ser Met Arg Pro Gln Pro Arg Val Leu Ser
                85                  90                  95
Gly Ala Pro Ser Leu Ala Ser Pro Gly His Thr Val Val Val Lys Thr
            100                 105                 110
Asp His Arg Gln Arg Leu Gln Cys Cys His Gly Phe Tyr Glu Ser Arg
        115                 120                 125
Gly Phe Cys Val Pro Leu Cys Ala Gln Glu Cys Val His Gly Arg Cys
    130                 135                 140
Val Ala Pro Asn Gln Cys Gln Cys Val Pro Gly Trp Arg Gly Asp Asp
145                 150                 155                 160
Cys Ser Ser Ala Pro Asn Cys Leu Gln Pro Cys Thr Pro Gly Tyr Tyr
                165                 170                 175
Gly Pro Ala Cys Gln Phe Arg Cys Gln Cys His Gly Ala Pro Cys Asp
            180                 185                 190
Pro Gln Thr Gly Ala Cys Phe Cys Pro Ala Glu Arg Thr Gly Pro Ser
        195                 200                 205
Cys Asp Val Ser Cys Ser Gln Gly Thr Ser Gly Phe Phe Cys Pro Ser
    210                 215                 220
Thr His Pro Cys Gln Asn Gly Gly Val Phe Gln Thr Pro Gln Gly Ser
```

```
225                 230                 235                 240
Cys Ser Cys Pro Pro Gly Trp Met Gly Thr Ile Cys Ser Leu Pro Cys
                245                 250                 255
Pro Glu Gly Phe His Gly Pro Asn Cys Ser Gln Glu Cys Arg Cys His
            260                 265                 270
Asn Gly Gly Leu Cys Asp Arg Phe Thr Gly Gln Cys Arg Cys Ala Pro
        275                 280                 285
Gly Tyr Thr Gly Asp Arg Cys Arg Glu Glu Cys Pro Val Gly Arg Phe
    290                 295                 300
Gly Gln Asp Cys Ala Glu Thr Cys Asp Cys Ala Pro Asp Ala Arg Cys
305                 310                 315                 320
Phe Pro Ala Asn Gly Ala Cys Leu Cys Glu His Gly Phe Thr Gly Asp
                325                 330                 335
Arg Cys Thr Asp Arg Leu Cys Pro Asp Gly Phe Tyr Gly Leu Ser Cys
            340                 345                 350
Gln Ala Pro Cys Thr Cys Asp Arg Glu His Ser Leu Ser Cys His Pro
        355                 360                 365
Met Asn Gly Glu Cys Ser Cys Leu Pro Gly Trp Ala Gly Leu His Cys
    370                 375                 380
Asn Glu Ser Cys Pro Gln Asp Thr His Gly Pro Gly Cys Gln Glu His
385                 390                 395                 400
Cys Leu Cys Leu His Gly Gly Val Cys Gln Ala Thr Ser Gly Leu Cys
                405                 410                 415
Gln Cys Ala Pro Gly Tyr Thr Gly Pro His Cys Ala Ser Leu Cys Pro
            420                 425                 430
Pro Asp Thr Tyr Gly Val Asn Cys Ser Ala Arg Cys Ser Cys Glu Asn
        435                 440                 445
Ala Ile Ala Cys Ser Pro Ile Asp Gly Glu Cys Val Cys Lys Glu Gly
    450                 455                 460
Trp Gln Arg Gly Asn Cys Ser Val Pro Cys Pro Pro Gly Thr Trp Gly
465                 470                 475                 480
Phe Ser Cys Asn Ala Ser Cys Gln Cys Ala His Glu Ala Val Cys Ser
                485                 490                 495
Pro Gln Thr Gly Ala Cys Thr Cys Thr Pro Gly Trp His Gly Ala His
            500                 505                 510
Cys Gln Leu Pro Cys Pro Lys Gly Gln Phe Gly Glu Gly Cys Ala Ser
        515                 520                 525
Arg Cys Asp Cys Asp His Ser Asp Gly Cys Asp Pro Val His Gly Arg
    530                 535                 540
Cys Gln Cys Gln Ala Gly Trp Met Gly Ala Arg Cys His Leu Ser Cys
545                 550                 555                 560
Pro Glu Gly Leu Trp Gly Val Asn Cys Ser Asn Thr Cys Thr Cys Lys
                565                 570                 575
Asn Gly Gly Thr Cys Leu Pro Glu Asn Gly Asn Cys Val Cys Ala Pro
            580                 585                 590
Gly Phe Arg Gly Pro Ser Cys Gln Arg Ser Cys Gln Pro Gly Arg Tyr
        595                 600                 605
Gly Lys Arg Cys Val Pro Cys Lys Cys Ala Asn His Ser Phe Cys His
    610                 615                 620
Pro Ser Asn Gly Thr Cys Tyr Cys Leu Ala Gly Trp Thr Gly Pro Asp
625                 630                 635                 640
Cys Ser Gln Pro Cys Pro Pro Gly His Trp Gly Glu Asn Cys Ala Gln
                645                 650                 655
Thr Cys Gln Cys His His Gly Gly Thr Cys His Pro Gln Asp Gly Ser
            660                 665                 670
Cys Ile Cys Pro Leu Gly Trp Thr Gly His His Cys Leu Glu Gly Cys
        675                 680                 685
Pro Leu Gly Thr Phe Gly Ala Asn Cys Ser Gln Pro Cys Gln Cys Gly
```

```
    690                 695                 700
Pro Gly Glu Lys Cys His Pro Glu Thr Gly Ala Cys Val Cys Pro Pro
705                 710                 715                 720
Gly His Ser Gly Ala Pro Cys Arg Ile Gly Ile Gln Glu Pro Phe Thr
                725                 730                 735
Val Met Pro Thr Thr Pro Val Ala Tyr Asn Ser
            740                 745
```

<210> 115
<211> 24
<212> PRT
<213> Homo sapiens

<400> 115

```
Leu Gly Ala Val Ile Gly Ile Ala Val Leu Gly Ser Leu Val Val Ala
 1               5                  10                  15
Leu Val Ala Leu Phe Ile Gly Tyr
            20
```

<210> 116
<211> 259
<212> PRT
<213> Homo sapiens

<400> 116

```
Arg His Trp Gln Lys Gly Lys Glu His His His Leu Ala Val Ala Tyr
 1               5                  10                  15
Ser Ser Gly Arg Leu Asp Gly Ser Glu Tyr Val Met Pro Asp Val Pro
            20                  25                  30
Pro Ser Tyr Ser His Tyr Tyr Ser Asn Pro Ser Tyr His Thr Leu Ser
        35                  40                  45
Gln Cys Ser Pro Asn Pro Pro Pro Pro Asn Lys Val Pro Gly Pro Leu
    50                  55                  60
Phe Ala Ser Leu Gln Asn Pro Glu Arg Pro Gly Gly Ala Gln Gly His
65                  70                  75                  80
Asp Asn His Thr Thr Leu Pro Ala Asp Trp Lys His Arg Arg Glu Pro
                85                  90                  95
Pro Pro Gly Pro Leu Asp Arg Gly Ser Ser Arg Leu Asp Arg Ser Tyr
            100                 105                 110
Ser Tyr Ser Tyr Ser Asn Gly Pro Gly Pro Phe Tyr Asp Lys Gly Leu
        115                 120                 125
Ile Ser Glu Glu Glu Leu Gly Ala Ser Val Ala Ser Leu Ser Ser Glu
    130                 135                 140
Asn Pro Tyr Ala Thr Ile Arg Asp Leu Pro Ser Leu Pro Gly Gly Pro
145                 150                 155                 160
Arg Glu Ser Ser Tyr Met Glu Met Lys Gly Pro Pro Ser Gly Ser Ala
                165                 170                 175
Pro Arg Gln Pro Pro Gln Phe Trp Asp Ser Gln Arg Arg Arg Gln Pro
            180                 185                 190
Gln Pro Gln Arg Asp Ser Gly Thr Tyr Glu Gln Pro Ser Pro Leu Ile
        195                 200                 205
His Asp Arg Asp Ser Val Gly Ser Gln Pro Pro Leu Pro Pro Gly Leu
    210                 215                 220
Pro Pro Gly His Tyr Asp Ser Pro Lys Asn Ser His Ile Pro Gly His
225                 230                 235                 240
Tyr Asp Leu Pro Pro Val Arg His Pro Pro Ser Pro Pro Leu Arg Arg
                245                 250                 255
Gln Asp Arg
```

<210> 117
<211> 497
<212> PRT
<213> Mus msuculus

<400> 117

```
Ser Thr His Ala Ser Gly Asp Pro Val His Gly Gln Cys Arg Cys Gln
 1               5                  10                  15
Ala Gly Trp Met Gly Thr Arg Cys His Leu Pro Cys Pro Glu Gly Phe
            20                  25                  30
Trp Gly Ala Asn Cys Ser Asn Thr Cys Thr Cys Lys Asn Gly Gly Thr
        35                  40                  45
Cys Val Ser Glu Asn Gly Asn Cys Val Cys Ala Pro Gly Phe Arg Gly
    50                  55                  60
Pro Ser Cys Gln Arg Pro Cys Pro Pro Gly Arg Tyr Gly Lys Arg Cys
65                  70                  75                  80
Val Gln Cys Lys Cys Asn Asn Asn His Ser Ser Cys His Pro Ser Asp
                85                  90                  95
Gly Thr Cys Ser Cys Leu Ala Gly Trp Thr Gly Pro Asp Cys Ser Glu
            100                 105                 110
Ala Cys Pro Pro Gly His Trp Gly Leu Lys Cys Ser Gln Leu Cys Gln
        115                 120                 125
Cys His His Gly Gly Thr Cys His Pro Gln Asp Gly Ser Cys Ile Cys
    130                 135                 140
Thr Pro Gly Trp Thr Gly Pro Asn Cys Leu Glu Gly Cys Pro Pro Arg
145                 150                 155                 160
Met Phe Gly Val Asn Cys Ser Gln Leu Cys Gln Cys Asp Leu Gly Glu
                165                 170                 175
Met Cys His Pro Glu Thr Gly Ala Cys Val Cys Pro Pro Gly His Ser
            180                 185                 190
Gly Ala Asp Cys Lys Met Gly Ser Gln Glu Ser Phe Thr Ile Met Pro
        195                 200                 205
Thr Ser Pro Val Thr His Asn Ser Leu Gly Ala Val Ile Gly Ile Ala
    210                 215                 220
Val Leu Gly Thr Leu Val Val Ala Leu Ile Ala Leu Phe Ile Gly Tyr
225                 230                 235                 240
Arg Gln Trp Gln Lys Gly Lys Glu His Glu His Leu Ala Val Ala Tyr
                245                 250                 255
Ser Thr Gly Arg Leu Asp Gly Ser Asp Tyr Val Met Pro Asp Val Ser
            260                 265                 270
Pro Ser Tyr Ser His Tyr Tyr Ser Asn Pro Ser Tyr His Thr Leu Ser
        275                 280                 285
Gln Cys Ser Pro Asn Pro Pro Pro Pro Asn Lys Val Pro Gly Ser Gln
    290                 295                 300
Leu Phe Val Ser Ser Gln Ala Pro Glu Arg Pro Ser Arg Ala His Gly
305                 310                 315                 320
Arg Glu Asn His Thr Thr Leu Pro Ala Asp Trp Lys His Arg Arg Glu
                325                 330                 335
Pro His Asp Arg Gly Ala Ser His Leu Asp Arg Ser Tyr Ser Cys Ser
            340                 345                 350
Tyr Ser His Arg Asn Gly Pro Gly Pro Phe Cys His Lys Gly Pro Ile
        355                 360                 365
Ser Glu Glu Gly Leu Gly Ala Ser Val Met Ser Leu Ser Ser Glu Asn
    370                 375                 380
Pro Tyr Ala Thr Ile Arg Asp Leu Pro Ser Leu Pro Gly Glu Pro Arg
385                 390                 395                 400
```

```
Glu Ser Gly Tyr Val Glu Met Lys Gly Pro Pro Ser Val Ser Pro Pro
                405                 410                 415
Arg Gln Ser Leu His Leu Arg Asp Arg Gln Gln Arg Gln Leu Gln Pro
            420                 425                 430
Gln Arg Asp Ser Gly Thr Tyr Glu Gln Pro Ser Pro Leu Ser His Asn
        435                 440                 445
Glu Glu Ser Leu Gly Ser Thr Pro Pro Leu Pro Pro Gly Leu Pro Pro
    450                 455                 460
Gly His Tyr Asp Ser Pro Lys Asn Ser His Ile Pro Gly His Tyr Asp
465                 470                 475                 480
Leu Pro Pro Val Arg His Pro Pro Ser Pro Pro Ser Arg Arg Gln Asp
                485                 490                 495
Arg
```

<210> 118
<211> 216
<212> PRT
<213> Mus musculus

<400> 118

```
Ser Thr His Ala Ser Gly Asp Pro Val His Gly Gln Cys Arg Cys Gln
 1               5                  10                  15
Ala Gly Trp Met Gly Thr Arg Cys His Leu Pro Cys Pro Glu Gly Phe
            20                  25                  30
Trp Gly Ala Asn Cys Ser Asn Thr Cys Thr Cys Lys Asn Gly Gly Thr
        35                  40                  45
Cys Val Ser Glu Asn Gly Asn Cys Val Cys Ala Pro Gly Phe Arg Gly
    50                  55                  60
Pro Ser Cys Gln Arg Pro Cys Pro Pro Gly Arg Tyr Gly Lys Arg Cys
65                  70                  75                  80
Val Gln Cys Lys Cys Asn Asn Asn His Ser Ser Cys His Pro Ser Asp
                85                  90                  95
Gly Thr Cys Ser Cys Leu Ala Gly Trp Thr Gly Pro Asp Cys Ser Glu
            100                 105                 110
Ala Cys Pro Pro Gly His Trp Gly Leu Lys Cys Ser Gln Leu Cys Gln
        115                 120                 125
Cys His His Gly Gly Thr Cys His Pro Gln Asp Gly Ser Cys Ile Cys
    130                 135                 140
Thr Pro Gly Trp Thr Gly Pro Asn Cys Leu Glu Gly Cys Pro Pro Arg
145                 150                 155                 160
Met Phe Gly Val Asn Cys Ser Gln Leu Cys Gln Cys Asp Leu Gly Glu
                165                 170                 175
Met Cys His Pro Glu Thr Gly Ala Cys Val Cys Pro Pro Gly His Ser
            180                 185                 190
Gly Ala Asp Cys Lys Met Gly Ser Gln Glu Ser Phe Thr Ile Met Pro
        195                 200                 205
Thr Ser Pro Val Thr His Asn Ser
    210                 215
```

<210> 119
<211> 24
<212> PRT
<213> Mus musculus

<400> 119

```
Leu Gly Ala Val Ile Gly Ile Ala Val Leu Gly Thr Leu Val Val Ala
 1               5                  10                  15
```

Leu Ile Ala Leu Phe Ile Gly Tyr
20

<210> 120
<211> 257
<212> PRT
<213> Mus musculus

<400> 120
Arg Gln Trp Gln Lys Gly Lys Glu His Glu His Leu Ala Val Ala Tyr
1 5 10 15
Ser Thr Gly Arg Leu Asp Gly Ser Asp Tyr Val Met Pro Asp Val Ser
20 25 30
Pro Ser Tyr Ser His Tyr Tyr Ser Asn Pro Ser Tyr His Thr Leu Ser
35 40 45
Gln Cys Ser Pro Asn Pro Pro Pro Pro Asn Lys Val Pro Gly Ser Gln
50 55 60
Leu Phe Val Ser Ser Gln Ala Pro Glu Arg Pro Ser Arg Ala His Gly
65 70 75 80
Arg Glu Asn His Thr Thr Leu Pro Ala Asp Trp Lys His Arg Arg Glu
85 90 95
Pro His Asp Arg Gly Ala Ser His Leu Asp Arg Ser Tyr Ser Cys Ser
100 105 110
Tyr Ser His Arg Asn Gly Pro Gly Pro Phe Cys His Lys Gly Pro Ile
115 120 125
Ser Glu Glu Gly Leu Gly Ala Ser Val Met Ser Leu Ser Ser Glu Asn
130 135 140
Pro Tyr Ala Thr Ile Arg Asp Leu Pro Ser Leu Pro Gly Glu Pro Arg
145 150 155 160
Glu Ser Gly Tyr Val Glu Met Lys Gly Pro Pro Ser Val Ser Pro Pro
165 170 175
Arg Gln Ser Leu His Leu Arg Asp Arg Gln Gln Arg Gln Leu Gln Pro
180 185 190
Gln Arg Asp Ser Gly Thr Tyr Glu Gln Pro Ser Pro Leu Ser His Asn
195 200 205
Glu Glu Ser Leu Gly Ser Thr Pro Pro Leu Pro Pro Gly Leu Pro Pro
210 215 220
Gly His Tyr Asp Ser Pro Lys Asn Ser His Ile Pro Gly His Tyr Asp
225 230 235 240
Leu Pro Pro Val Arg His Pro Pro Ser Pro Pro Ser Arg Arg Gln Asp
245 250 255
Arg

<210> 121
<211> 636
<212> PRT
<213> Rattus sp.

<400> 121
Met Gly Val Ile Cys Ser Leu Pro Cys Pro Glu Gly Phe His Gly Pro
1 5 10 15
Asn Cys Thr Gln Glu Cys Arg Cys His Asn Gly Gly Leu Cys Asp Arg
20 25 30
Phe Thr Gly Gln Cys His Cys Ala Pro Gly Tyr Ile Gly Asp Arg Cys
35 40 45
Arg Glu Glu Cys Pro Val Gly Arg Phe Gly Gln Asp Cys Ala Glu Thr
50 55 60

```
Cys Asp Cys Ala Pro Gly Ala Arg Cys Phe Pro Ala Asn Gly Ala Cys
65                  70                  75                  80
Leu Cys Glu His Gly Phe Thr Gly Asp Arg Cys Thr Glu Arg Leu Cys
                85                  90                  95
Pro Asp Gly Arg Tyr Gly Leu Ser Cys Gln Asp Pro Cys Thr Cys Asp
            100                 105                 110
Pro Glu His Ser Leu Ser Cys His Pro Met His Gly Glu Cys Ser Cys
        115                 120                 125
Gln Pro Gly Trp Ala Gly Leu His Cys Asn Glu Ser Cys Pro Gln Asp
    130                 135                 140
Thr His Gly Ala Gly Cys Gln Glu His Cys Leu Cys Leu His Gly Gly
145                 150                 155                 160
Val Cys Leu Ala Asp Ser Gly Leu Cys Arg Cys Ala Pro Gly Tyr Thr
                165                 170                 175
Gly Pro His Cys Ala Asn Leu Cys Pro Pro Asn Thr Tyr Gly Ile Asn
            180                 185                 190
Cys Ser Ser His Cys Ser Cys Glu Asn Ala Ile Ala Cys Ser Pro Val
        195                 200                 205
Asp Gly Thr Cys Ile Cys Lys Glu Gly Trp Gln Arg Gly Asn Cys Ser
    210                 215                 220
Val Pro Cys Pro Pro Gly Thr Trp Gly Phe Ser Cys Asn Ala Ser Cys
225                 230                 235                 240
Gln Cys Ala His Glu Gly Val Cys Ser Pro Gln Thr Gly Ala Cys Thr
                245                 250                 255
Cys Thr Pro Gly Trp Arg Gly Val His Cys Gln Leu Pro Cys Pro Lys
            260                 265                 270
Gly Gln Phe Gly Glu Gly Cys Ala Ser Val Cys Asp Cys Asp His Ser
        275                 280                 285
Asp Gly Cys Asp Pro Val His Gly His Cys Arg Cys Gln Ala Gly Trp
    290                 295                 300
Met Gly Thr Arg Cys His Leu Pro Cys Pro Glu Gly Phe Trp Gly Ala
305                 310                 315                 320
Asn Cys Ser Asn Ala Cys Thr Cys Lys Asn Gly Gly Thr Cys Val Pro
                325                 330                 335
Glu Asn Gly Asn Cys Val Cys Ala Pro Gly Phe Arg Gly Pro Ser Cys
            340                 345                 350
Gln Arg Pro Cys Pro Pro Gly Arg Tyr Gly Lys Arg Cys Val Pro Cys
        355                 360                 365
Lys Cys Asn Asn His Ser Ser Cys His Pro Ser Asp Gly Thr Cys Ser
    370                 375                 380
Cys Leu Ala Gly Trp Thr Gly Pro Asp Cys Ser Glu Ser Cys Pro Pro
385                 390                 395                 400
Gly His Trp Gly Leu Lys Cys Ser Gln Pro Cys Gln Cys His His Gly
                405                 410                 415
Ala Thr Cys His Pro Gln Asp Gly Ser Cys Val Cys Ile Pro Gly Trp
            420                 425                 430
Thr Gly Pro Asn Cys Ser Glu Gly Cys Pro Ser Arg Met Phe Gly Val
        435                 440                 445
Asn Cys Ser Gln Leu Cys Gln Cys Asp Pro Gly Glu Met Cys His Pro
    450                 455                 460
Glu Thr Gly Ala Cys Val Cys Pro Pro Gly His Ser Gly Ala His Cys
465                 470                 475                 480
Lys Val Gly Ser Gln Glu Ser Phe Thr Ile Met Pro Thr Ser Pro Val
                485                 490                 495
Ile His Asn Ser Leu Gly Ala Val Ile Gly Ile Ala Val Leu Gly Thr
            500                 505                 510
Leu Val Val Ala Leu Val Ala Leu Phe Ile Gly Tyr Arg His Trp Gln
        515                 520                 525
```

```
Lys Gly Lys Glu His Glu His Leu Ala Val Ala Tyr Ser Thr Gly Arg
    530                 535                 540
Leu Asp Gly Ser Asp Tyr Val Met Pro Asp Val Ser Pro Ser Tyr Ser
545                 550                 555                 560
His Tyr Tyr Ser Asn Pro Ser Tyr His Thr Leu Ser Gln Cys Ser Pro
                565                 570                 575
Asn Pro Pro Pro Pro Asn Lys Ile Pro Gly Ser Gln Leu Phe Val Ser
            580                 585                 590
Ser Gln Ala Ser Glu Arg Pro Asn Arg Asn His Gly Arg Asp Asn His
        595                 600                 605
Ala Thr Leu Pro Ala Asp Trp Lys His Arg Arg Glu Ser His Asp Arg
    610                 615                 620
Ala Phe Leu Arg His Gln Pro Pro Gly Pro Lys Val
625                 630                 635
```

<210> 122
<211> 500
<212> PRT
<213> Rattus sp.

<400> 122

```
Met Gly Val Ile Cys Ser Leu Pro Cys Pro Glu Gly Phe His Gly Pro
 1               5                  10                  15
Asn Cys Thr Gln Glu Cys Arg Cys His Asn Gly Gly Leu Cys Asp Arg
            20                  25                  30
Phe Thr Gly Gln Cys His Cys Ala Pro Gly Tyr Ile Gly Asp Arg Cys
        35                  40                  45
Arg Glu Glu Cys Pro Val Gly Arg Phe Gly Gln Asp Cys Ala Glu Thr
    50                  55                  60
Cys Asp Cys Ala Pro Gly Ala Arg Cys Phe Pro Ala Asn Gly Ala Cys
65                  70                  75                  80
Leu Cys Glu His Gly Phe Thr Gly Asp Arg Cys Thr Glu Arg Leu Cys
                85                  90                  95
Pro Asp Gly Arg Tyr Gly Leu Ser Cys Gln Asp Pro Cys Thr Cys Asp
            100                 105                 110
Pro Glu His Ser Leu Ser Cys His Pro Met His Gly Glu Cys Ser Cys
        115                 120                 125
Gln Pro Gly Trp Ala Gly Leu His Cys Asn Glu Ser Cys Pro Gln Asp
    130                 135                 140
Thr His Gly Ala Gly Cys Gln Glu His Cys Leu Cys Leu His Gly Gly
145                 150                 155                 160
Val Cys Leu Ala Asp Ser Gly Leu Cys Arg Cys Ala Pro Gly Tyr Thr
                165                 170                 175
Gly Pro His Cys Ala Asn Leu Cys Pro Pro Asn Thr Tyr Gly Ile Asn
            180                 185                 190
Cys Ser Ser His Cys Ser Cys Glu Asn Ala Ile Ala Cys Ser Pro Val
        195                 200                 205
Asp Gly Thr Cys Ile Cys Lys Glu Gly Trp Gln Arg Gly Asn Cys Ser
    210                 215                 220
Val Pro Cys Pro Pro Gly Thr Trp Gly Phe Ser Cys Asn Ala Ser Cys
225                 230                 235                 240
Gln Cys Ala His Glu Gly Val Cys Ser Pro Gln Thr Gly Ala Cys Thr
                245                 250                 255
Cys Thr Pro Gly Trp Arg Gly Val His Cys Gln Leu Pro Cys Pro Lys
            260                 265                 270
Gly Gln Phe Gly Glu Gly Cys Ala Ser Val Cys Asp Cys Asp His Ser
        275                 280                 285
Asp Gly Cys Asp Pro Val His Gly His Cys Arg Cys Gln Ala Gly Trp
```

```
    290                 295                 300
Met Gly Thr Arg Cys His Leu Pro Cys Pro Glu Gly Phe Trp Gly Ala
305                 310                 315                 320
Asn Cys Ser Asn Ala Cys Thr Cys Lys Asn Gly Gly Thr Cys Val Pro
                325                 330                 335
Glu Asn Gly Asn Cys Val Cys Ala Pro Gly Phe Arg Gly Pro Ser Cys
            340                 345                 350
Gln Arg Pro Cys Pro Pro Gly Arg Tyr Gly Lys Arg Cys Val Pro Cys
        355                 360                 365
Lys Cys Asn Asn His Ser Ser Cys His Pro Ser Asp Gly Thr Cys Ser
    370                 375                 380
Cys Leu Ala Gly Trp Thr Gly Pro Asp Cys Ser Glu Ser Cys Pro Pro
385                 390                 395                 400
Gly His Trp Gly Leu Lys Cys Ser Gln Pro Cys Gln Cys His His Gly
                405                 410                 415
Ala Thr Cys His Pro Gln Asp Gly Ser Cys Val Cys Ile Pro Gly Trp
            420                 425                 430
Thr Gly Pro Asn Cys Ser Glu Gly Cys Pro Ser Arg Met Phe Gly Val
        435                 440                 445
Asn Cys Ser Gln Leu Cys Gln Cys Asp Pro Gly Glu Met Cys His Pro
    450                 455                 460
Glu Thr Gly Ala Cys Val Cys Pro Pro Gly His Ser Gly Ala His Cys
465                 470                 475                 480
Lys Val Gly Ser Gln Glu Ser Phe Thr Ile Met Pro Thr Ser Pro Val
                485                 490                 495
Ile His Asn Ser
            500

      <210> 123
      <211> 24
      <212> PRT
      <213> Rattus sp.

      <400> 123
Leu Gly Ala Val Ile Gly Ile Ala Val Leu Gly Thr Leu Val Val Ala
 1               5                  10                  15
Leu Val Ala Leu Phe Ile Gly Tyr
            20

      <210> 124
      <211> 112
      <212> PRT
      <213> Rattus sp.

      <400> 124
Arg His Trp Gln Lys Gly Lys Glu His Glu His Leu Ala Val Ala Tyr
 1               5                  10                  15
Ser Thr Gly Arg Leu Asp Gly Ser Asp Tyr Val Met Pro Asp Val Ser
            20                  25                  30
Pro Ser Tyr Ser His Tyr Tyr Ser Asn Pro Ser Tyr His Thr Leu Ser
        35                  40                  45
Gln Cys Ser Pro Asn Pro Pro Pro Pro Asn Lys Ile Pro Gly Ser Gln
    50                  55                  60
Leu Phe Val Ser Ser Gln Ala Ser Glu Arg Pro Asn Arg Asn His Gly
65                  70                  75                  80
Arg Asp Asn His Ala Thr Leu Pro Ala Asp Trp Lys His Arg Arg Glu
                85                  90                  95
Ser His Asp Arg Ala Phe Leu Arg His Gln Pro Pro Gly Pro Lys Val
```

```
            100                 105                 110

      <210> 125
      <211> 28
      <212> PRT
      <213> Homo sapiens

      <400> 125
Met Ala Pro Ala Arg Ala Gly Phe Cys Pro Leu Leu Leu Leu Leu Leu
 1               5                  10                  15
Leu Gly Leu Trp Val Ala Glu Ile Pro Val Ser Ala
            20                  25

      <210> 126
      <211> 128
      <212> PRT
      <213> Homo sapiens

      <400> 126
Lys Pro Lys Gly Met Thr Ser Ser Gln Trp Phe Lys Ile Gln His Met
 1               5                  10                  15
Gln Pro Ser Pro Gln Ala Cys Asn Ser Ala Met Lys Asn Ile Asn Lys
            20                  25                  30
His Thr Lys Arg Cys Lys Asp Leu Asn Thr Phe Leu His Glu Pro Phe
        35                  40                  45
Ser Ser Val Ala Ala Thr Cys Gln Thr Pro Lys Ile Ala Cys Lys Asn
    50                  55                  60
Gly Asp Lys Asn Cys His Gln Ser His Gly Pro Val Ser Leu Thr Met
65                  70                  75                  80
Cys Lys Leu Thr Ser Gly Lys Tyr Pro Asn Cys Arg Tyr Lys Glu Lys
                85                  90                  95
Arg Gln Asn Lys Ser Tyr Val Val Ala Cys Lys Pro Pro Gln Lys Lys
            100                 105                 110
Asp Ser Gln Gln Phe His Leu Val Pro Val His Leu Asp Arg Val Leu
        115                 120                 125

      <210> 127
      <211> 19
      <212> PRT
      <213> Homo sapiens

      <400> 127
Met Pro Leu Leu Thr Leu Tyr Leu Leu Leu Phe Trp Leu Ser Gly Tyr
 1               5                  10                  15
Ser Ile Ala


      <210> 128
      <211> 286
      <212> PRT
      <213> Homo sapiens

      <400> 128
Thr Gln Ile Thr Gly Pro Thr Thr Val Asn Gly Leu Glu Arg Gly Ser
 1               5                  10                  15
Leu Thr Val Gln Cys Val Tyr Arg Ser Gly Trp Glu Thr Tyr Leu Lys
            20                  25                  30
Trp Trp Cys Arg Gly Ala Ile Trp Arg Asp Cys Lys Ile Leu Val Lys
```

```
        35                  40                  45
Thr Ser Gly Ser Glu Gln Glu Val Lys Arg Asp Arg Val Ser Ile Lys
    50                  55                  60
Asp Asn Gln Lys Asn Arg Thr Phe Thr Val Thr Met Glu Asp Leu Met
65                  70                  75                  80
Lys Thr Asp Ala Asp Thr Tyr Trp Cys Gly Ile Glu Lys Thr Gly Asn
                85                  90                  95
Asp Leu Gly Val Thr Val Gln Val Thr Ile Asp Pro Ala Ser Thr Pro
            100                 105                 110
Ala Pro Thr Thr Pro Thr Ser Thr Thr Phe Thr Ala Pro Val Thr Gln
        115                 120                 125
Glu Glu Thr Ser Ser Ser Pro Thr Leu Thr Gly His His Leu Asp Asn
    130                 135                 140
Arg His Lys Leu Leu Lys Leu Ser Val Leu Leu Pro Leu Ile Phe Thr
145                 150                 155                 160
Ile Leu Leu Leu Leu Leu Val Ala Ala Ser Leu Leu Ala Trp Arg Met
                165                 170                 175
Met Lys Tyr Gln Gln Lys Ala Ala Gly Met Ser Pro Glu Gln Val Leu
            180                 185                 190
Gln Pro Leu Glu Gly Asp Leu Cys Tyr Ala Asp Leu Thr Leu Gln Leu
        195                 200                 205
Ala Gly Thr Ser Pro Arg Lys Ala Thr Thr Lys Leu Ser Ser Ala Gln
    210                 215                 220
Val Asp Gln Val Glu Val Glu Tyr Val Thr Met Ala Ser Leu Pro Lys
225                 230                 235                 240
Glu Asp Ile Ser Tyr Ala Ser Leu Thr Leu Gly Ala Glu Asp Gln Glu
                245                 250                 255
Pro Thr Tyr Cys Asn Met Gly His Leu Ser Ser His Leu Pro Gly Arg
            260                 265                 270
Gly Pro Glu Glu Pro Thr Glu Tyr Ser Thr Ile Ser Arg Pro
        275                 280                 285

      <210> 129
      <211> 150
      <212> PRT
      <213> Homo sapiens

      <400> 129
Thr Gln Ile Thr Gly Pro Thr Thr Val Asn Gly Leu Glu Arg Gly Ser
 1               5                  10                  15
Leu Thr Val Gln Cys Val Tyr Arg Ser Gly Trp Glu Thr Tyr Leu Lys
            20                  25                  30
Trp Trp Cys Arg Gly Ala Ile Trp Arg Asp Cys Lys Ile Leu Val Lys
        35                  40                  45
Thr Ser Gly Ser Glu Gln Glu Val Lys Arg Asp Arg Val Ser Ile Lys
    50                  55                  60
Asp Asn Gln Lys Asn Arg Thr Phe Thr Val Thr Met Glu Asp Leu Met
65                  70                  75                  80
Lys Thr Asp Ala Asp Thr Tyr Trp Cys Gly Ile Glu Lys Thr Gly Asn
                85                  90                  95
Asp Leu Gly Val Thr Val Gln Val Thr Ile Asp Pro Ala Ser Thr Pro
            100                 105                 110
Ala Pro Thr Thr Pro Thr Ser Thr Thr Phe Thr Ala Pro Val Thr Gln
        115                 120                 125
Glu Glu Thr Ser Ser Ser Pro Thr Leu Thr Gly His His Leu Asp Asn
    130                 135                 140
Arg His Lys Leu Leu Lys
145                 150
```

<210> 130
<211> 24
<212> PRT
<213> Homo sapiens

<400> 130
Leu Ser Val Leu Leu Pro Leu Ile Phe Thr Ile Leu Leu Leu Leu Leu
1 5 10 15
Val Ala Ala Ser Leu Leu Ala Trp
20

<210> 131
<211> 112
<212> PRT
<213> Homo sapiens

<400> 131
Arg Met Met Lys Tyr Gln Gln Lys Ala Ala Gly Met Ser Pro Glu Gln
1 5 10 15
Val Leu Gln Pro Leu Glu Gly Asp Leu Cys Tyr Ala Asp Leu Thr Leu
20 25 30
Gln Leu Ala Gly Thr Ser Pro Arg Lys Ala Thr Thr Lys Leu Ser Ser
35 40 45
Ala Gln Val Asp Gln Val Glu Val Glu Tyr Val Thr Met Ala Ser Leu
50 55 60
Pro Lys Glu Asp Ile Ser Tyr Ala Ser Leu Thr Leu Gly Ala Glu Asp
65 70 75 80
Gln Glu Pro Thr Tyr Cys Asn Met Gly His Leu Ser Ser His Leu Pro
85 90 95
Gly Arg Gly Pro Glu Glu Pro Thr Glu Tyr Ser Thr Ile Ser Arg Pro
100 105 110

<210> 132
<211> 21
<212> PRT
<213> Homo sapiens

<400> 132
Met Asp His Cys Gly Ala Leu Phe Leu Cys Leu Cys Leu Leu Thr Leu
1 5 10 15
Gln Asn Ala Thr Thr
20

<210> 133
<211> 507
<212> PRT
<213> Homo sapiens

<400> 133
Glu Thr Trp Glu Glu Leu Leu Ser Tyr Met Glu Asn Met Gln Val Ser
1 5 10 15
Arg Gly Arg Ser Ser Val Phe Ser Ser Arg Gln Leu His Gln Leu Glu
20 25 30
Gln Met Leu Leu Asn Thr Ser Phe Pro Gly Tyr Asn Leu Thr Leu Gln
35 40 45
Thr Pro Thr Ile Gln Ser Leu Ala Phe Lys Leu Ser Cys Asp Phe Ser
50 55 60
Gly Leu Ser Leu Thr Ser Ala Thr Leu Lys Arg Val Pro Gln Ala Gly

```
65                  70                  75                  80
Gly Gln His Ala Arg Gly Gln His Ala Met Gln Phe Pro Ala Glu Leu
                85                  90                  95
Thr Arg Asp Ala Cys Lys Thr Arg Pro Arg Glu Leu Arg Leu Ile Cys
            100                 105                 110
Ile Tyr Phe Ser Asn Thr His Phe Phe Lys Asp Glu Asn Asn Ser Ser
        115                 120                 125
Leu Leu Asn Asn Tyr Val Leu Gly Ala Gln Leu Ser His Gly His Val
    130                 135                 140
Asn Asn Leu Arg Asp Pro Val Asn Ile Ser Phe Trp His Asn Gln Ser
145                 150                 155                 160
Leu Glu Gly Tyr Thr Leu Thr Cys Val Phe Trp Lys Glu Gly Ala Arg
                165                 170                 175
Lys Gln Pro Trp Gly Gly Trp Ser Pro Glu Gly Cys Arg Thr Glu Gln
            180                 185                 190
Pro Ser His Ser Gln Val Leu Cys Arg Cys Asn His Leu Thr Tyr Phe
        195                 200                 205
Ala Val Leu Met Gln Leu Ser Pro Ala Leu Val Pro Ala Glu Leu Leu
    210                 215                 220
Ala Pro Leu Thr Tyr Ile Ser Leu Val Gly Cys Ser Ile Ser Ile Val
225                 230                 235                 240
Ala Ser Leu Ile Thr Val Leu Leu His Phe His Phe Arg Lys Gln Ser
                245                 250                 255
Asp Ser Leu Thr Arg Ile His Met Asn Leu His Ala Ser Val Leu Leu
            260                 265                 270
Leu Asn Ile Ala Phe Leu Leu Ser Pro Ala Phe Ala Met Ser Pro Val
        275                 280                 285
Pro Gly Ser Ala Cys Thr Ala Leu Ala Ala Ala Leu His Tyr Ala Leu
    290                 295                 300
Leu Ser Cys Leu Thr Trp Met Ala Ile Glu Gly Phe Asn Leu Tyr Leu
305                 310                 315                 320
Leu Leu Gly Arg Val Tyr Asn Ile Tyr Ile Arg Arg Tyr Val Phe Lys
                325                 330                 335
Leu Gly Val Leu Gly Trp Gly Ala Pro Ala Leu Leu Val Leu Leu Ser
            340                 345                 350
Leu Ser Val Lys Ser Ser Val Tyr Gly Pro Cys Thr Ile Pro Val Phe
        355                 360                 365
Asp Ser Trp Glu Asn Gly Thr Gly Phe Gln Asn Met Ser Ile Cys Trp
    370                 375                 380
Val Arg Ser Pro Val Val His Ser Val Leu Val Met Gly Tyr Gly Gly
385                 390                 395                 400
Leu Thr Ser Leu Phe Asn Leu Val Val Leu Ala Trp Ala Leu Trp Thr
                405                 410                 415
Leu Arg Arg Leu Arg Glu Arg Ala Asp Ala Pro Ser Val Arg Ala Cys
            420                 425                 430
His Asp Thr Val Thr Val Leu Gly Leu Thr Val Leu Leu Gly Thr Thr
        435                 440                 445
Trp Ala Leu Ala Phe Phe Ser Phe Gly Val Phe Leu Leu Pro Gln Leu
    450                 455                 460
Phe Leu Phe Thr Ile Leu Asn Ser Leu Tyr Gly Phe Phe Leu Phe Leu
465                 470                 475                 480
Trp Phe Cys Ser Gln Arg Cys Arg Ser Glu Ala Glu Ala Lys Ala Gln
                485                 490                 495
Ile Glu Ala Phe Ser Ser Ser Gln Thr Thr Gln
            500                 505
```

<210> 134
<211> 223

<212> PRT
<213> Homo sapiens

<400> 134
Glu Thr Trp Glu Glu Leu Leu Ser Tyr Met Glu Asn Met Gln Val Ser
1 5 10 15
Arg Gly Arg Ser Ser Val Phe Ser Ser Arg Gln Leu His Gln Leu Glu
20 25 30
Gln Met Leu Leu Asn Thr Ser Phe Pro Gly Tyr Asn Leu Thr Leu Gln
35 40 45
Thr Pro Thr Ile Gln Ser Leu Ala Phe Lys Leu Ser Cys Asp Phe Ser
50 55 60
Gly Leu Ser Leu Thr Ser Ala Thr Leu Lys Arg Val Pro Gln Ala Gly
65 70 75 80
Gly Gln His Ala Arg Gly Gln His Ala Met Gln Phe Pro Ala Glu Leu
85 90 95
Thr Arg Asp Ala Cys Lys Thr Arg Pro Arg Glu Leu Arg Leu Ile Cys
100 105 110
Ile Tyr Phe Ser Asn Thr His Phe Phe Lys Asp Glu Asn Asn Ser Ser
115 120 125
Leu Leu Asn Asn Tyr Val Leu Gly Ala Gln Leu Ser His Gly His Val
130 135 140
Asn Asn Leu Arg Asp Pro Val Asn Ile Ser Phe Trp His Asn Gln Ser
145 150 155 160
Leu Glu Gly Tyr Thr Leu Thr Cys Val Phe Trp Lys Glu Gly Ala Arg
165 170 175
Lys Gln Pro Trp Gly Gly Trp Ser Pro Glu Gly Cys Arg Thr Glu Gln
180 185 190
Pro Ser His Ser Gln Val Leu Cys Arg Cys Asn His Leu Thr Tyr Phe
195 200 205
Ala Val Leu Met Gln Leu Ser Pro Ala Leu Val Pro Ala Glu Leu
210 215 220

<210> 135
<211> 25
<212> PRT
<213> Homo sapiens

<400> 135
Leu Ala Pro Leu Thr Tyr Ile Ser Leu Val Gly Cys Ser Ile Ser Ile
1 5 10 15
Val Ala Ser Leu Ile Thr Val Leu Leu
20 25

<210> 136
<211> 20
<212> PRT
<213> Homo sapiens

<400> 136
Leu His Ala Ser Val Leu Leu Leu Asn Ile Ala Phe Leu Leu Ser Pro
1 5 10 15
Ala Phe Ala Met
20

<210> 137
<211> 21
<212> PRT

<213> Homo sapiens

<400> 137
Tyr Ala Leu Leu Ser Cys Leu Thr Trp Met Ala Ile Glu Gly Phe Asn
1 5 10 15
Leu Tyr Leu Leu Leu
20

<210> 138
<211> 19
<212> PRT
<213> Homo sapiens

<400> 138
Leu Gly Val Leu Gly Trp Gly Ala Pro Ala Leu Leu Val Leu Leu Ser
1 5 10 15
Leu Ser Val

<210> 139
<211> 25
<212> PRT
<213> Homo sapiens

<400> 139
Val Leu Val Met Gly Tyr Gly Gly Leu Thr Ser Leu Phe Asn Leu Val
1 5 10 15
Val Leu Ala Trp Ala Leu Trp Thr Leu
20 25

<210> 140
<211> 21
<212> PRT
<213> Homo sapiens

<400> 140
Val Thr Val Leu Gly Leu Thr Val Leu Leu Gly Thr Thr Trp Ala Leu
1 5 10 15
Ala Phe Phe Ser Phe
20

<210> 141
<211> 20
<212> PRT
<213> Homo sapiens

<400> 141
Leu Phe Leu Phe Thr Ile Leu Asn Ser Leu Tyr Gly Phe Phe Leu Phe
1 5 10 15
Leu Trp Phe Cys
20

<210> 142
<211> 24
<212> PRT
<213> Homo sapiens

<400> 142

```
Ser Gln Arg Cys Arg Ser Glu Ala Glu Ala Lys Ala Gln Ile Glu Ala
 1               5                  10                  15
Phe Ser Ser Ser Gln Thr Thr Gln
            20
```

<210> 143
<211> 16
<212> PRT
<213> Homo sapiens

<400> 143

```
Ser Pro Val Pro Gly Ser Ala Cys Thr Ala Leu Ala Ala Ala Leu His
 1               5                  10                  15
```

<210> 144
<211> 37
<212> PRT
<213> Homo sapiens

<400> 144

```
Lys Ser Ser Val Tyr Gly Pro Cys Thr Ile Pro Val Phe Asp Ser Trp
 1               5                  10                  15
Glu Asn Gly Thr Gly Phe Gln Asn Met Ser Ile Cys Trp Val Arg Ser
            20                  25                  30
Pro Val Val His Ser
        35
```

<210> 145
<211> 7
<212> PRT
<213> Homo sapiens

<400> 145

```
Gly Val Phe Leu Leu Pro Gln
 1               5
```

<210> 146
<211> 17
<212> PRT
<213> Homo sapiens

<400> 146

```
His Phe His Phe Arg Lys Gln Ser Asp Ser Leu Thr Arg Ile His Met
 1               5                  10                  15
Asn
```

<210> 147
<211> 14
<212> PRT
<213> Homo sapiens

<400> 147

```
Gly Arg Val Tyr Asn Ile Tyr Ile Arg Arg Tyr Val Phe Lys
 1               5                  10
```

<210> 148
<211> 18

<212> PRT
<213> Homo sapiens

<400> 148

```
Arg Arg Leu Arg Glu Arg Ala Asp Ala Pro Ser Val Arg Ala Cys His
 1               5                  10                  15
Asp Thr
```

A. CLASSIFICATION OF SUBJECT MATTER

IPC(7) :C07K 14/47; C07H 21/04; C12N 15/63; C12P 21/02

US CL : 530/350; 536/23.5; 435/320.1, 252.3, 361, 69.1

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 530/350; 536/23.5; 435/320.1, 252.3, 361, 69.1

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

Commercial Sequence Databases: GenEmbl, EST, Issued_Patents_NA, N_Geneseq_36, PIR_64, SwissProt_38, A_Geneseq_36,Issued_Patents_AA, SPTREMBL_12

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	Database EST, AN AQ588144, ZHOU et al. 'CITBI-E1-2644L24.TF CITBI-B1 Homo sapiens genomic clone 2644L24, genomic survey sequence'. 07 June 1999, see attached alignment showing 100% identical match to nucleotides 88-481 of SEQ ID NO: 1 (394 nucleotides total).	1, 3, 5
Y		2, 4, 6-10 and 12
A	Database SPTREMBL_12, AN Q28396, RICHARDSON et al. 'Type II Collagen from Equus caballus (Horse)'. 01 November 1996. Polypeptide 25.7% identical to the amino acid sequence of SEQ ID NO:2, see attached alignment, Nov. 1, 1996.	1-10 and 12

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

*	Special categories of cited documents:	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A"	document defining the general state of the art which is not considered to be of particular relevance	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E"	earlier document published on or after the international filing date	"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L"	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&"	document member of the same patent family
"O"	document referring to an oral disclosure, use, exhibition or other means		
"P"	document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search 21 SEPTEMBER 2000	Date of mailing of the international search report 02 OCT 2000
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230	Authorized officer EILEEN B. O'HARA Telephone No. (703) 308-0196

Form PCT/ISA/210 (second sheet) (July 1998)*

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ **Claims Nos.:**
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ **Claims Nos.:**
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ **Claims Nos.:**
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

Please See Extra Sheet.

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:
1-10 and 12

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest.

☐ No protest accompanied the payment of additional search fees.

BOX II. OBSERVATIONS WHERE UNITY OF INVENTION WAS LACKING
This ISA found multiple inventions as follows:

This application contains the following inventions or groups of inventions which are not so linked as to form a single inventive concept under PCT Rule 13.1. In order for all inventions to be searched, the appropriate additional search fees must be paid.

Group I, claim(s)1-10 and 12, in so far as they are drawn to Intercept 340, polynucleotides of SEQ ID NOS: 1 and 3, vector, host cell, method of producing a protein recombinantly and protein of SEQ ID NO: 2.
Groups II-VII, claim(s) 1-10 and 12, in so far as they are drawn to the next six polynucleotides of distinct cDNA clones and encoded proteins, identified as Mango 003, Mango 347, Tango 272, Tango 295, Tango 354 and Tango 378, as listed in Tables 1 and 2.
Groups VIII-XIV, claim(s) 11 and 15, in so far as they are drawn to antibodies to one of the seven proteins listed above.
Groups XV-XXI, claims 13, 14, 19, 20 and 22, in so far as they are drawn to a method for detecting the presence of in a sample or identifying a compound which binds to or modulates the activity of a polypeptide of one of the seven proteins listed above.
Groups XXII-XXVII, claims 16 and 17, in so far as they are drawn to a method for detecting the nucleic acids of one of the seven cDNA clones listed above.
Groups XXIX-XXXV, claim 18, in so far as it is drawn to a kit comprising a compound of unspecified constitution which selectively binds to a nucleic acid molecule of the seven cDNA clones listed above.
Groups XXXVI-XLII, claim 21, in so far as it is drawn to a method for modulating the activity of one of the seven proteins listed above.

The inventions listed as Groups I-XLII do not relate to a single inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons: Group I corresponds to the first invention wherein the first product is the polynucleotide and the first method of using is the method of making the protein. Note that there is no method of making the polynucleotide. The invention also includes the protein made. Each of groups II-VII does not share the same or corresponding special technical feature because each group is drawn to a different polynucleotide and encoded protein, and each of groups VIII-XLII does not share the same or corresponding special technical feature because each group is drawn to different compounds or methods of using the seven polynucleotides and encoded proteins. This Authority therefore considers that the several inventions do not share a special technical feature within the meaning of PCT Rule 13.2 and thus do not relate to a single general inventive concept within the meaning of PCT Rule 13.1.